

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 03:38:42 ; Search time 6582.52 Seconds
(without alignments)
840.674 Million cell updates/sec

Title: US-09-300-482-14

Perfect score: 410
Sequence: 1 cccacgcgtccgcgcatg.....gcaacgcccgtcttcgatg 410

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estcpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estcl:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pin:*
16: em_gss_vrl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	346.8	84.6	530	10	BE363077	DGL_9_F06
2	343.8	83.9	492	10	BI141708	IP1_55_H1
3	323.8	79.0	531	10	BE267183	HV_CEA001
4	323.8	79.0	590	9	AV939826	AV939826
5	323.8	79.0	612	9	AL504224	AL504224
6	323.8	79.0	722	9	BE216705	HV_CEB001
7	323.8	79.0	886	10	BE413061	MCG015_E0
8	322.2	78.6	445	10	BM370860	EBR004_SO
9	321.4	78.4	833	10	BI951719	HVSMEM000
10	320.6	78.2	466	10	BE471136	BE471136
11	319	77.8	531	10	BI543081	WHE0284_C
12	319	77.8	547	10	BE493869	WHE1276_F
13	319	77.8	823	9	AM448282	BRY_1785
14	317	77.3	875	10	BF630844	HVSMEM001
15	315.6	77.0	536	10	BI074740	IP1_15_F0
16	314.6	76.7	519	9	AM536648	LGL_245_A
17	313.6	76.5	516	10	BE363025	DGL_9_A10

18	312	76.1	438	10	BG053648 RH122_8_A
19	310.6	75.8	547	9	AV938902 AV938902
20	309	75.4	539	9	AV924281 AV924281
21	307.4	75.0	466	10	BF483000 WHE2313_E
22	307.4	75.0	667	10	BI954565 HVSMEM001
23	295.2	72.0	743	10	BE414610 SCUD011_F0
24	294.4	71.8	473	10	BE657125 OV2_21_B0
25	290.4	70.8	456	10	D24970 RICR2869A_R
26	290.2	70.8	472	10	BE497862 WHE0957_D
27	290.2	70.8	474	10	BE606368 WHE0904_B
28	288.6	70.4	420	10	BF201489 WHE1771_B
29	288.6	70.4	474	10	BE497880 WHE0957_B
30	288.6	70.4	512	10	BE425231 WHE0312_G
31	284	69.3	788	10	BI947260 HVSMEM100
32	283.8	69.2	338	10	BG948605 IP1_11_G0
33	281.4	68.6	454	10	BE495944 WHE1259_B
34	281.2	68.6	456	10	BI074623 IP1_13_C0
35	277.4	67.7	781	10	BE414643 SCUD011_H1
36	277	67.6	427	9	AU233121 AU233121
37	274	66.8	453	10	BM326433 PIC1_57_A
38	270.6	66.0	446	10	BG262764 WHE0945_A
39	269	65.6	463	9	AM286879 LGL_222_F
40	268.4	65.5	683	10	BG370009 HVSMEM1002
41	268	65.4	462	9	AM286839 LGL_222_B
42	267	65.1	461	9	AM286858 LGL_222_D
43	266	64.9	460	9	AM677090 DGL_4_B11
44	266	64.9	468	10	BG947868 IP1_6_C11
45	265.8	64.8	467	10	BG053295 RH122_25_

ALIGNMENTS

RESULT 1
LOCUS BE363077 530 bp mRNA linear EST 20-JUL-2000
DEFINITION DGL_9_F06.b1_A002 Dark Grown 1 (DGL) Sorghum bicolor CDNA, mRNA
sequence.
ACCESSION BE363077
VERSION BE363077.1 GI:9304634
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 530)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt ,L.H.
An EST database from Sorghum: dark-grown seedlings
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmp@pratt.uga.edu
Sequences have been trimmed to exclude polyA, vector and regions below phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: JEN REV
High quality sequence stop: 453
POLYA-No.

FEATURES

source
Location/Qualifiers
1..530
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DGL)"
/note="Organ: 5-day-old dark-grown seedlings; Vector: Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."

BASE COUNT ORIGIN	126 a	132 c	164 g	108 t
----------------------	-------	-------	-------	-------

Query Match	84.6%;	Score 346.8;	DB 10;	Length 530;
Best Local Similarity	91.7%;	Pred. No. 8.3e-70;		
Matches 366; Conservative	0;	Mismatches 33;	Indels 0;	Gaps 0;

QY 12 gcgycatcgtgggcagaaacctgcgccttaacattgcagagaaggcttcccatctcgtg 71
Db 91 GCGGTCATGAGGGGAGAACCTTGCCCTCAATATTGCAGAAAGGCTTCCCATCTCTGTC 150
QY 72 tacaaacaggaacaaacctccaaggctgacagagaccgtgcagcgtgcgaagcagaaagaaac 131
Db 151 TACAACAGAGAACCTTCAAGGTTGATGAGACCGTGCACGCTGCCAAGCAGAAAGGAAC 210
QY 132 ctcccgctcctacggtcttcacatgaccccggtctcttgaaatccattccagaagccaag 191
Db 211 CTTCCTCTGTACGGCTTCCATGCACCTCGATCCTTTGTGTAATCTCATTAACAAACCCCGT 270
QY 192 gtggtgacatcgtcgtcacaagccgagccgagccagttgacacagaccatcgagagctgca 251
Db 271 GTTCTCATCATGCTGTGCACGCTGGTGACACAGTTGAACACATATGCAGCCCTTGCA 330
QY 252 gctactctggaagagagcgacatgcatacgcgatgggggagaaagatggtacgaagaacag 311
Db 331 GCTACTCTGAGACAGGGGTGACTGTATACATATGATGAGAGGAAATGAGTATAGAACAACT 390
QY 312 gaagaagaggaagaagccatgagagagcgccgcctcncgtatctctgagatgagtgctctc 371
Db 391 GAGAGGAGGAGAGGCAATGGAGAGAGCCTGGCCCTCTTATCTTTGGAAATGGGCGTCTT 450
QY 372 gtagagaaagagaggtgcccgcacaagggccgctctctgata 410
Db 451 GGAGGAGAGGAGGCTCCCGCAACAGACCATCTTGTAG 489

RESULT	2
BII41708	
LOCUS	
DEFINITION	BII41708 492 bp mRNA linear EST 03-JUL-2001 IPI_55_H11.b1_A002 Immature panicle 1 (IPI) sorghum bicolor cDNA, mRNA sequence.
ACCESSION	BII41708
VERSION	BII41708.1 GI:14594151
KEYWORDS	EST,
SOURCE	sorghum.
ORGANISM	sorghum bicolor
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum. Klein,R.R., Cordomier-Pratt,M.-M., Gingle,A., Sudman,M. and Pratt L.H. An EST database from Sorghum: developing preanthesis panicles Unpublished (2001) Contact: Cordomier-Pratt MM
TITLE	
JOURNAL	
COMMENT	

```

FEATURES
    source
    Location/Qualifiers
    1..492
    /organism="Sorghum bicolor"
    /cultivar="BTx623"

```

```

/db_xref="taxon:4558"
/clone_lib="Immature panning 1 (IP1)"
/notes="Organ: Developing preanthesis panicles; Vector:
pBluescript II SK(-) from Lambda zap II; Site_1: XhoI;
Site_2: EcoRI; The library was made from poly-A RNA in the
cloning vector lambda zap II. Clones to be sequenced were
prepared by mass excision."
BASE COUNT      115 a      125 c      154 g      98 t
ORIGIN

```

Query Match	83.9%	Score 343.8;	DB 10;	Length 492;
Best Local Similarity	91.7%;	Pred. No. 4e-69;		
Matches 363; Conservative	0;	Mismatches 33;	Indels 0;	Gaps 0;

Query Match	83.9%	Score 343.8;	DB 10;	Length 492;
Best Local Similarity	91.7%;	Pred. No. 4e-69;		
Matches 363; Conservative	0;	Mismatches 33;	Indels 0;	Gaps 0;

QY	12	gcggtaatgggacagaccttgcgcctcaacatttgcagagaagggttcccatctctgtg	71
Db	97	gcggatattggggacgaaccttgccttcaacatttgacagaaaggctttcccatctctgtc	156
QY	72	tacaacagagcaaacctcccaaggctgacagagacgtgtcacgctgtccaaagcagaaggaac	131
Db	157	TACAACAGGCAACCTCCAAAGTTGATGAGACCGTCAGCTGCCAAGGCAAGAAAGAAC	216
QY	132	cttcgccctcaagaccttcatacccgcgctcttggaaagtccattcagaagccaag	191
Db	217	CTTCCTGTGTACGGCTTCATGACCTTCATCTTGTGAACTTCATTCAGAAAGCCGCT	276
QY	192	gttggtatcaatgctcgtcaaggccgagccgacagctgaacagaccatcgcagcctcgca	251
Db	277	GTTGTGATCATGCTTGTCMAAGGCTGGTGCACAGTTGACAGAACATNTGGAGCTTCGA	336
QY	252	gtccactctggagcagggcgactgcatactgaatgggggaaagatggtatcagagaacag	311
Db	337	GCCTCATTTGGACAGGGGTACTGTCTATTATGAAAGGATAGTGGATGAGAACT	396
QY	312	gagaaggaaggagaagcattgaggaagcgcgcctnctgactcttgacatggtgtcct	371
Db	397	GAAAGGAGGGAAGAACCATGGAGAGGAGTGCCCTCTCTATCTTGGAATGGCGCTCT	456
QY	372	ggaggaagaagggtgtcccgcaagcgccgctccttg	407
Db	457	GGAGGAGAGGAGGTGCCCCGCAAGGAACTCACTTGG	492

RESULT	3				
LOCUS	BF267183				
DEFINITION	BF267183	531 bp	mRNA	linear	EST 23-OCT-2001
	HV.CE20017D02f Hordeum vulgare seedling green leaf				EST library
	HV.DN0004 (Blumeria challenged)				Hordeum vulgare cDNA clone
ACCESSION	HV.CE20017D02f				mRNA sequence.
VERSION	BF267183				
KEYWORDS	BF267183.2	GI:13263227			
SOURCE	EST.				
ORGANISM	barley.				
	Hordeum vulgare				

Unpublished (2001)
On Nov 17, 2000 this sequence version replaced g1:1198178.
Contact: Wang RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293

FEATURES	source
DB	429 TGGAGGAGAGGAGGTGCCCCGATGGCCGTCATGATG 468
RESULT	5
AL504224	
LOCUS	
DEFINITION	AL504224 612 bp mRNA linear EST 04-JAN-2001
ACCESSION	AL504224 Hordeum vulgare Barke roots Hordeum vulgare cDNA clone
VERSION	AL504224
KEYWORDS	AL504224.1 GI:12030439
SOURCE	EST.
ORGANISM	barley.
REFERENCE	Hordeum vulgare
AUTHORS	EunaxiVota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae
JOURNAL	; Triticeae; Hordeum.
COMMENT	1 (bases 1 to 612)
	Michaelak.W., Weschke,W., Pleissner,K.-P. and Graner,A.
	EST sequencing and analysis in barley
	Unpublished (2000)
	Contact: Michaelak.W
	Institute for Plant Genetics and Crop Plant Research
	Corrensstr.3, D-06466 Gatersleben, Germany
	Email: michaelak@p.kgatersleben.de, http://pgrc.ipk-gatersleben.de
	Seq primer: T3 primer for 5'end.
	Location/Qualifiers
	1..612

BASE COUNT	ORIGIN	ORIGIN	ORIGIN
140 a	166 c	177 g	126 t
uniquely. The maximum length was set to 700 bp			
3 others			

Query Match	79.0%;	Score 323.8;	DB 9;	Length 612;
Best Local Similarity	88.0%;	Pred. No. 1.7e-64;		
Matches 352; Conservative	0;	Mismatches 48;	Indels 0;	Gaps 0;

[illegible]

Accession	Sequence	Position
Db	AGCAACCTGGAGCGGCGCACTGCATCGTTGACGGAGAAACAGTCGATATGAAAC	346
Oy	ggaagagagaggaagagagcatatgagagagagcgagcctctatcttgcattgcatggtgtctc	370
Db	AGAAAGGAGCGGAGAAAGCGATGAGAGAGACGTGGACCTCTCTAAGCTCGATGGGTGTTTC	406
Oy	tggagagaaagtaggtgtccgcgaacagggcccgctctcatg	410
Db	TGGAGAGAGGAGGATGCCGCCCATGGCCCGTCATCATG	446

RESULT	6	
LOCUS	BE216705	
DEFINITION	BE216705	722 bp mRNA linear EST 23-OCT-2001
ACCESSION	HY_CEB001116f	Hordeum vulgare seedling green leaf EST library
VERSION	HY_CEB001116f.1	Plumeria challenged Hordeum vulgare cDNA clone
KEYWORDS	BE216705	mRNA sequence.
SOURCE	BE216705.2	GI:13265056
ORGANISM	barley.	
	Hordeum vulgare	

REFERENCE
AUTHORS
Wing, R., Close, T.J., Kleinholz, A., Wise, R., Wei, F., Baum, D.,
1 (bases 1 to 722)
Eukaryota; viridiplantae; streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Trillaceae; Hordeum.
1 (bases 1 to 722)

TITLE
"W., Fenton, R.D., Oates, R. and Main, D.
Development of a genetically anchored EST resource
for barley genomics: Blumeria Infected Incompatible (MIa6) seedling
leaf cDNA library
Unpublished (2001)
On Jul 3, 2000 this sequence version replaced gl:8904317.
8

FEATURES
 source
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwinge@clemson.edu
 Total bp bases = 459
 Seq primer: ATTATACCTTCCTACTAAGG
 High quality sequence stop: 617.
 Location/Qualifiers
 1..722

```

/organism="Hordeum vulgare"
/cultivar="C116151 (Mla6)"
/db_xref="taxon:4513"
/clone="HY_CED0011Et6f"
/clone_lib="Hordeum vulgare seedling green leaf EST
library HVCDNA0005 (Blumeria challenged)"
/tissue_type="seedling green leaf"
/lab_host="SOLR"
/notes="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
C.I. 16151 (Mla6) plants were greenhouse grown in the R
wise lab at Iowa State University, Ames, IA; 7 day old
green seedlings were challenged with isolate 5874 (AVRMA6
of Blumeria graminis f. sp. hordei) and leaves were
harvested 20 and 24 hr post-inoculation and snap frozen;
uninoculated leaves were harvested 20 hr post-inoculation
(Wei, Wise). In the T4 Close lab at the University of
California, Riverside, total RNA was prepared from each
sample pool, equal quantities of all three RNA pools were
combined, poly(A) RNA was purified from the mixture, one
primary unamplified cDNA library was made, and 1 million
pfu were in vivo excised to give plusescript SK(-) cDNA
phageids (Choi, Close). Phageids were plated and picked
at the Clemson University Genomics Institute (CUGI) (Begum
, Palmer, Frisch, Atkins and Wing). Plasmid DNA
preparations, DNA sequencing and sequence analysis were
performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates
, Ranbo, Main). The sequence has been trimmed to remove

```


vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wang R, Kleinofn A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/g93pages/bqng/31/Cover.html>)"

Query Match	79.0%;	Score 323.8;	DB 9;	Length 722;
Best Local Similarity	88.0%;	Pred. No. 1.8e-64;		
Matches 352; Conservative	0;	Mismatches 48;	Indels 0;	Gaps 0

OY	11	cgcgctcatgggagaaaccttgcctcaacattcaagaaaggtttcccatctctt	70
Db	51	CGCGCTCATTGGGGCAGAAACCTTGCCCTCAACATTGCAGAAAGGCTTCCCATCTCTGT	110
OY	71	gtacacagagacaacctccaagtgtagcagagacgltgacagcgtgcaccaagcagaagaa	130
Db	111	GTACACAGAGACACATCCAAAGTGTGAGCAGACCGTCACGCCGCCCAACTTGAAGAAA	170
OY	131	cccttcggtctacggcttccatlgaccccgcgctccttgtgaatccatcagaagccagc	190
Db	171	CCTTCCCTCTTACGGGTGTTTCCATGACCCCTGCACTCTTTGTCACTCCATTCAGAGCCAGC	230
OY	191	ggtgtgtgtcatgctcgttcaagcgcgcgcgccagttgaaccaacatctgagagctgc	250
Db	231	TGTGCTCATCATGCTGTGTCAAGGCCGGTCTCCGTGTGACAGACCATTGCGCACCGCTCCG	290
OY	251	agctcaacttgaagcagcgagctgtcatctcogatgggggaacagttgttacgaagaac	310
Db	291	AGCNAACCTGGAGCAGCGGGGACTGCATCTTTACGAGAGAAACGATGGTATGAGAACAC	350
OY	311	ggagagagaggaagccatgtggaagagcgcggcctnctgtactctgtgcatagtgtctc	370
Db	351	AGAAAGGAGGGGAGAGGGCATGTGAGAGAGCGTGGACTCTCACTTCGTGATGGGTGTTTC	410
OY	371	tggaggaagagaggtgtcccgcaagcgcccgcttcttgatg	410
Db	411	TGGAGGAGAGGAGGTGTCCCGCATGGCCCGCTTCATGATG	450

RESULT	7			
BE13061				
LOCUS				
DEFINITION	BE13061	886 bp	mRNA	linear
ACCESSION	MC0015.E05R990625	ITEC	MC6 Barley Leaf/Culm Library	EST 24-JUL-2000
VERSION	CDNA clone	MC0015.E05	mRNA sequence.	
KEYWORDS	BE13061			
SOURCE	BE13061.1	GI:9410907		
ORGANISM	EST.			
	Barley.			
	<i>Hordeum vulgare</i>			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae			
REFERENCE	1 (bases 1 to 886)			
AUTHORS	Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier			
	,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,			
	Herrmann,R.G., Holton,T., Jacquemund,J.M., Jia,J., Joudrier,P.,			
	Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,			
	Pechioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,			
	Sorrells,M., Wadburton,M. and Wenzel,G.			
TITLE	International Triticeae EST Cooperative (ITEC): Production of			
JOURNAL	Expressed Sequence Tags for Species of the Triticeae			
COMMENT	Unpublished (2000)			
	Contact: Graner A			
	Institute for Plant Genetics & Crop Plant Research			
	Cottrenstr. 3, D-06466 Gatersleben GERMANY			
	Tel: 49 39482 5521			

Fax: 49 39482 5137
Email: a_graner@ipk-gatersleben.de
International Triflicae EST Cooperative (ITEC)
<http://wheat.pw.usda.gov/gsnome>
<http://www.iti.uni-erlangen.de>

BASE COUNT	202 a	224 c	258 g	186 t	16 others
ORIGIN					

	Query Match	Best Local Similarity	79.0%;	Score 323.8;	DB 10;	Length 886;
	Matches	352;	Conservative	0;	Mismatches	48;
					Indels	0;
					Gaps	0;
OY	11	cgcgcatggtggcgcgaacacttgccttcaacatgtgcagagaaaggttcccatctctgt	70			
Db	47	CGCGCTCATGTGGGGCGAGAACCTTGGCCCTCAACATTGCCAGAAAGGCTTCCCATCTTCTGT	106			
OY	71	gtacaacggggaacactctccaagggtgagagaaacgtgagagcgttgcccaaggcgaagagaa	130			
Db	107	CTACACACGAGCGACATCCAAAGTGTGAGCAGACCTCCAGCGCCGCAAGCTTGAAAGGAAA	166			
OY	131	ccctcccatctacaggtcttcacatgaccccgcttcccttltgaaagtcacatcagaagccag	190			
Db	167	CTTCTCCTCTCAAGGTTTTCATGACCCGTCAGATCCTTTGTCAACCTCATTTACGAAGCCACG	226			
OY	191	ggtgtgtcatcatgtctgcgtcaaggccgagcgccgaattgaccagacatcgcgaagctgc	250			
Db	227	TGTCTCTCATCATGCTTGTTCAGGCGCGGTCTCCGGTTTACACAGACATTCGGCACCTCTGC	286			
OY	251	agctcaacttggagcagggcgactgcatacatcgatgtggggagacagagttgtgtcagaaacac	310			
Db	287	AGCACACCTTGGAGCGAGGGCGACTGCATCGTTGACGCGAAGAAACGAGTGTATGAACAC	346			
OY	311	ggaagagagaggaagagccatgtgagagcgcgcctnctgtatcttggcatgggtgtctc	370			
Db	347	AGAAAGGAGGAGAAAGCGATGTGAGAGACGTCGATCCTTCAACCTCGGTATGGGTGTTC	406			
OY	371	tggagagaaaggaggtgcccgcgaagggccgttccattgt	410			
Db	407	TGGAGAGAGAGGGGTGCCCGCATGGCGCCGCTGCATGTATG	446			

RESULT	8
LOCUS	BM370860
DEFINITION	445 bp mRNA linear EST 10-JAN-2002 EBR004_S0002_J06_R IGF Barley EBR004 library Hordeum vulgare cDNA clone EBR004_S0002_J06 5', mRNA sequence.
ACCESSION	BM370860
VERSION	BM370860.1
KEYWORDS	GI:18114250
SOURCE	EST.
ORGANISM	barley. Hordeum vulgare
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae; Triticeae; Hordeum.
AUTHORS	1 (bases 1 to 445) Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L., Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.
TITLE	Development of Barley Transcriptome Resources
JOURNAL	Unpublished (2001)
COMMENT	Contact: Waugh R Unit of Genomics Scottish Crop Research Institute Invergowrie, Dundee, DD2 5DA, Scotland, UK Tel: 00 44 1382 562731 Fax: 00 44 1382 562426

Db	182	GCATCATGCTTGTCAAGGCGGTGCTCGGTTCACCAAGACCATGCGCCACCGCTCCAGCA	241
QY	255	cacttggagcagggcgacgtctcatcgtatgggggaaacgagtggtacgagaacaacgag	314
Db	242	CACCTGGAGCAGGGGCGACATGCTATCGTACGAGGAGAAACGATGGTATGAGAACACAGA	301
QY	315	aggagggagaagggcattggagagcgcgccnctgtactcttgatgtgggtgtctctgga	374
Db	302	AGGAGGGAGAAAGCCATGGAGAGGGTGGACCTCTACTCGGTATGGGTGTTCTTGA	361
QY	375	ggaagagaggtgtccgcacgcccgcgtcttattc	410
Db	362	GGAGAGGAGGTTGCCCGCCCATGCGCCCTTCATGATG	397

RESULT	10
BE471136	
LOCUS	BE471136
DEFINITION	466 bp mRNA linear EST 28-JUL-2000
WHE0284_C07_E14S	wheat drought-stressed seedling cDNA library
Triticum aestivum	cDNA clone WHE0284_C07_E14, mRNA sequence.
BE471136	
ACCESSION	BE471136

ACCESSION	BE471130
VERSION	BE471136.1
KEYWORDS	GI:9561627
SOURCE	EST.
ORGANISM	bread wheat. <i>Triticum aestivum</i>

REFERENCE
AUTHORS

1 (bases 1 to 466)
Anderson, O.D., Chao, S., Chol, D.W., Close, T.J., Fenton, R.D., Han

TITLE The structure and function of the expressed portion of the wheat genomes - Drought-stressed seedling cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: coanders@nrcp.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: StrataGene SK primer.

FEATURES	Location/Qualifiers
SOURCE	1, .466

```

/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE0284_C07_E14"
/clone_lib="Wheat drought-stressed seedling cdna library"
/tissue_type="Seedling without endosperm"
/dev_stage="Five day Old seedling"

```

/ra₁-pscl-E. coli SQR
/note-Vector: Lambda Uni-ZAP XR, excised phagemid;
Site-1: EcoRI; Site-2: XhoI; Seeds were surface-sterilized,
germinated and grown aseptically in the dark at room
temperature on filter paper with water, nystatin and
cefotaxime in covered crystallization dishes. Five-day old
seedlings were incubated for one day at 90% RH. After
removing endosperm, seedlings were transferred to

desiccator jar containing saturated MgSO₄ at room temperature for 24 hr. The tissue, total RNA and poly(A)⁺ RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give phagescript phagemids in the TJ Clonase lab (Choi, Clonase, Fentron) at the University of California, Riverside. plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors). *

BASE COUNT	106 a	134 c	135 g	91 t
ORIGIN				

Query Match	78.2%	Score 320.6;	DB 10;	Length 466;
Best Local Similarity	87.5%;	Pred. No. 8.3e-64;		
Matches 350; Conservative	0;	Mismatches 50;	Indels 0;	Gaps 0;

11	cgcggtatcgggcagaaacctctgctccctcaacattgcagagaagaaggtctcccatctcgt	70
12		
13		
14		
49	CGCGGTATCGGGCCAGAACCTTGCCCTCAACATTGCAAGAAAGGGTTGCCCTCTCTGT	108
50		
71	gtataaacaagacaacccccaagttgtagcagagaccgtgcagctgycgaagcgagaagaa	130
72		
109	CTACAAACAGGACCACTCCCAAGGTCGATGAGACTGTTACACGGCCCAACTCAAAAGGAAA	168
110		
131	ccttcgcgctacagcgtctccatgaccgccggtcctctgtgtaagttccattcaagaagcacg	190
132		
169	CCTTCCTCTCTACGGTTTCATGACCCCTGCATCTCCTTGCTGTAATCTCATTCAGAAACCAAG	228
170		
191	ggtgtgatatcatctcgttcaagccgagcgagcgaattgacacagacatcgagcgtctgc	250
192		
229	TGTCGTATCATGCTTGTCAAGGCGGCTGCTCCTGTTGACCAACACCATGCAACGCTTCC	288
230		
251	agctcactcttgagagagcgagactgcatactcagtcagtgagggaagcaagtgtacagaagac	310
252		
289	AGCACACTGAGACAGGCGCGACTGCTCATCTGTTGATGGAGCAAAAGAGTGGTACAGAAACAC	348
290		
311	ggaagaggagaggaagagccatlgagaagacgcgagcctctatctctgacatggagtgctc	370
312		
349	GGAAGAGAGGAGCAAGCGATGAGGAGCGCTGGGACCTCTTACTCTCGGAGTGGGTGTTTC	408
350		
371	tggaggaagagaggtgtcccgcaacgagccggtctcttgatg	410
372		
409	CGGAGGAGAGGAGGTTCCCGCCACTGAGCCCTCCATCATGATG	448

	RESULT	11
	LOCUS	B1543081
DEFINITION	B1543081 531 bp mRNA linear EST 04-SEP-2001	
ACCESSION	9490702HD.V1 949 - Juvenile leaf and shoot CDNA from Steve Moose	
VERSION	B1543081.1 GI:15427259	

ELEMENTS
 SOURCE
 ORGANISM
 EST.
 Zee may.
 Zee may.
 Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zee.
 1 (bases 1 to 531)
 Walhot.V

TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL	Unpublished (1999)
COMMENT	Contact: Walbot V

Department of Biological Sciences
Stanford University
855 California Ave., Palo Alto, CA 94304, USA
Tel.: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 949072.row: H.column: 04.
row+column+plate

FEATURES	LOCATION/VALUES
SOURCE	1. .531

```

/organism="Zea mays"
/cultivar="W64A"
/db_xref="taxon:4577"
/clone_lib="949 - Juvenile leaf and shoot cDNA from Steve
Moose"
/tissue_type="Immature leaf primordium and vegetative
meristem"
/dev_stage="4 stages from 3-13 days after imbibing"

```

```

/lab_host="E. coli XL0LR"
/note="Organ: juvenile vegetative shoots; Vector:
PAD-GAL4-2.1; Site_1: EcoRI; Site_2: XhoI; Equal amounts

```


of total RNA by weight from 4 tissue sources (see below) were pooled, polyA⁺ RNA isolated, and cDNA synthesized for EcoRI (5') and XhoI (3') directional cloning into lambda HybriZap vector from Stratagene. Tissue Sources: 1. Whole shoots 3 days after sowing/imbibing in wet soil. 2. Basal 1.5 cm shoots 6 days after sowing - includes yellow portions of developing leaves 1-5, primordia from 6-8, and the vegetative apex. 3. Non-green portions of developing leaves 4-5 and the vegetative apex, including adult leaf primordia, 9 days after sowing. 4. Partially expanded and green leaves 4-5 at 13 days after sowing."

BASE COUNT 105 a 156 c 173 g 97 t

ORIGIN

Query Match 77.8%; Score 319; DB 10; Length 531;
Best Local Similarity 87.2%; Pred. No. 2e-63;

Matches 349; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 11 cgcggtcatgggagagaaccttgcctcaacattgacagagaagggtcccatctctgt 70
DB 103 CGCGGTCTATGGAGACACTTGCCTCAACATCGCGGAGAAAGGTTCCCTATCTCGGT 162
QY 71 gtacacagagacaacctccaaggtgagacagcgttgacgctgccaagagagaa 130
DB 163 CTACACAGACAGACCTCCAGAGTTGATGAGACGTCGACAGCTCCAGAGTCAAGGAA 222
QY 131 ccttcgcgtctacgcttccatgacccgcgctccttcttgtagtccattcagagccagc 190
DB 223 CCTCCCGCTGTGGTTGTCACGACCCGCGCTCTGTGAGTCATCCAGAACCCCG 282
QY 191 ggttgtagtctgctgtcgaagcgcgcgccagttgacacagacatcgacgctcgc 250
DB 283 TGTGTCATCTATGCTCGTCAAGGCTGGGCGCCGCTGACACAGACCATTTGCGACGCTCGC 342
QY 251 agctcaacttgagagagcgagctgcatcatcgatgagggagagagagtgtagcagaaac 310
DB 343 GCGCACCTTGATCGAGGGGAGCTGATGCTGATGTCGACAGAGTGATGAGAACAC 402
QY 311 ggaagagagagagagagcattgagagagcgcgcccttctgtatcttgagcatgggtctc 370
DB 403 GGAAGAGAGGAGAGGAGCGATGAGGAGGAGCGGCGCTCTTATCTTGTCATGGCGCTCTC 462
QY 371 tggaggaagagaggtgcccgcgaacgcccgtccttgatg 410
DB 463 CGGAGGAGAGGAGGCTGCCCGCATGGCCCTCTGATG 502

RESULT 12

BE493869

LOCUS

DEFINITION WHE1276_F08_K16ZS Secale cereale anther cDNA library Secale cereale

ACCESSION BE493869

VERSION BE493869.1 GI:9660462

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Anderson, O.D., Butler, E., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Gustafson, J.P., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Ross, K., Seaton, C.L. and Tong, J.C.
The structure and function of the expressed portion of the wheat genomes - Another cDNA library from rye

UNPUBLISHED (2000)

CONTACT: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific

West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105959773

Fax: 510595818
Email: oanderson@w.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
Location/Qualifiers
1. 547
/organism="Secale cereale"
/cultivar="Blanco"
/db_xref="taxon:4550"
/clone="WHE1276_F08_K16"
/clone_1lb="Secale cereale anther cDNA library"
/tissue_type="Anther"
/dev_stage="Adult plant before anthesis"
/lab_host="E. coli S04R"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown in the greenhouse. Anthers were harvested and pooled from early meiosis to late meiosis. The tissue, total RNA, and poly(A) RNA were prepared (Butler, Ross and Gustafson) at University of Missouri, Columbia. A cDNA library was made, and the cDNA clones were in vivo excised to give phagescript phagemids in the TJ Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT 121 a 149 c 160 g 117 t

ORIGIN

Query Match 77.8%; Score 319; DB 10; Length 547;
Best Local Similarity 87.2%; Pred. No. 2.1e-63;

Matches 349; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 11 cgcggtcatgggagagaaccttgcctcaacattgacagagaagggtcccatctctgt 70
DB 48 CGCGGTCTATGGAGACACTTGCCTCAACATTCGCGGAGAAAGGTTCCCTATCTCGGT 107
QY 71 gtacacagagacaacctccaaggtgagacagcgttgacgctgccaagagagaa 130
DB 108 CTACACAGACAGACCTCCAGAGTTGATGAGACGTCGACAGCTCCAGAGTCAAGGAA 167
QY 131 ccttcgcgtctacgcttccatgacccgcgctccttcttgtagtccattcagagccagc 190
DB 168 CCTTCCTCTATGCTCGTCAAGGCTGGGCGCCGCTGACACAGACCATTTGCGACGCTCGC 227
QY 191 ggttgtagtctgctgtcgaagcgcgcgccagttgacacagacatcgacgctcgc 250
DB 228 CCGGTCTATGCTCGTCAAGGCTGGGCGCTGTCGACAGACCATTCGCGACGCTCGC 287
QY 251 agctcaacttgagagagcgagctgcatcatcgatgagggagagagtgtagcagaaac 310
DB 288 TGCACACCTGAGAGCAAGTGACTGCTGTTGATGAGAGAAACGATGATGAGAACAC 347
QY 311 ggaagagagagagagcattgagagagcgcgcccttctgtatcttgagcatgggtctc 370
DB 348 GGAAGAGAGGAGAGGCGATGAGAGCGTGTGACTCTTCATCTCGGAGATGGGTTC 407
QY 371 tggaggaagagaggtgcccgcgaacgcccgtccttgatg 410
DB 408 CGGAGGAGAGGAGGCGCCCGCATGGCCCTCTGATG 447

RESULT 13

AM448282

LOCUS

DEFINITION BR1785 BRY Trilicium aestivum cDNA clone P61-10, mRNA sequence.

ACCESSION AM448282

VERSION AM448282.1 GI:12018817

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae			
AUTHORS	1 (bases 1 to 823)			
TITLE	Triticaceae; Triticum.			
JOURNAL	Clarke/B.C., Hobbs/M. and Appels,R. Genes active in developing wheat endosperm Unpublished (2000)			
COMMENT	Contact: Bryan Clarke Division of Plant Industry C.S.I.R.O. GPO Box 1600, Canberra, ACT, Australia Tel: 61 2 6246 5054 Fax: 61 2 6246 5000 Email: bryan@cpi.csiro.au.			
FEATURES	location/Qualifiers			
source	1..823			
	/organism="Triticum aestivum"			
	/cultivar="Wynna"			
	/db_xref="taxon:4565"			
	/clone="P61-10"			
	/clone_1fb="BRY"			
	/cell_type="endosperm"			
BASE COUNT	180 a 233 c 247 g 163 t			
ORIGIN				
Query Match	77.8%; Score 319; DB 9; Length 823;			
Best Local Similarity	87.2%; Pred.No.2.4e-63;			
Matches 349; Conservative 0; Mismatches 51; Indels 0; Gaps 0;				
OY	11	CGCGGTCATGGGCGAGAACCTTGCCTCAATATGCAAGAAAGGTTCCCATCTCTGT	70	
DB	153	CGCGCTCATGGGGCAAGAACCTTGCCCTCAACCTTGCGAGAAAGGCTTCCCATCTCGT	212	
OY	71	GTACAACAGGACCAACTCAAGTGTGACGAGACGCTGACGCTGCAAGCAAGAGAA	130	
DB	213	CTACAAACAGGACCACTCCCAAGGTGAGAGACGCTTACGCGCCCAAGCTGAGAGAAA	272	
OY	131	CCCTCCGCTCAAGGCTCCATGATACCCCGTCTCTTGTGAAGTCCATCAGAGCACG	190	
DB	273	CCCTCCCTCTCAACGGTTTCCATGACCCCTGCATCCCTCGTCAACTTCATCAAGCCACG	332	
OY	191	GTGTGTGATCATGTCGTCGAAGCCGAGCGCCAGTTACAGACCATCGAGCTCGC	250	
DB	333	TGTGCTCATTCATGCTTGTTCMAAGCCGGCGGCTTCGGTTTACACGACTATCGCACCTCGC	392	
OY	251	AGCTCACTTGAAGCAAGGCGAGCTGATCATCGATGTGGGAGAGATGTAAGAGAACAC	310	
DB	393	AGCAACACTGGAGCAGGCTGACTGCATGCTGATGAGAGAAACGAGTGCTACAGAACAC	452	
OY	311	GGAGAGGAGGAGAAAGCCATGGAGAGACGCGGCTTCTATCTTGCATGGAGTGTCTC	370	
DB	453	GCAAAAGGAGGAGAAAGCGATGAGAGACGTGACTCTTCACTCGGAGATGGGTTC	512	
OY	371	TGAAGAGAAAGAGGTGCGCGCAAGCGCCGCTCTGATG	410	
DB	513	CGGAGAGAGAGAGGTGCCCGCCATGCGCCCTCATGATG	552	
RESULT 14				
BF630844	875 bp mRNA linear EST 22-OCT-2001			
LOCUS	BF630844			
DEFINITION	HVSME0014D11f Hordeum vulgare seedling shoot EST library			
	HVCND00002 (dehydration stress) Hordeum vulgare cDNA clone			
	HVSME0014D11f, mRNA sequence.			
ACCESSION	BF630844			
VERSION	BF630844.2 GI:13091746			
KEYWORDS	EST.			
SOURCE	barley.			
ORGANISM	Hordeum vulgare			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae			
	; Triticaceae; Hordeum.			
REFERENCE	1 (bases 1 to 875)			

FEATURES	source
JOURNAL COMMENT	Unpublished (2001) On Dec 19, 2000 this sequence version replaced g1:11895002. Contact: Wing RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 7288 Fax: 864 656 4293 Email: rwing@clemson.edu Total hg bases = 444 Seq primer: AATTACCTTCACTAAGG High quality sequence stop: 543. Location/Qualifiers 1. 875
FEATURES	<pre> /organism="Hordeum vulgare" /cultivar="Morex" /db_xref="taxon:4513" /clone="HYSMED0014D11f" /clone_id="Hordeum vulgare seedling shoot EST library HVCdNA0002 (Dehydration stress)" /tissue_type="Seedling shoot" /lab_host="TJC121" /notes="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI; Seeds were surface sterilized then germinated under axenic conditions in the dark at room temperature on filter paper with water, nystatin and ceftioaxime in covered crystallization dishes. Five-day old seedlings were incubated at 90% RH for 24 hr. Shoots were then harvested, total RNA was prepared, poly(A) RNA was purified, one primary unamplified cDNA library was made, 600000 pfu were in vivo excised to give plasmid SK(-) cDNA phagemids. These steps were performed in the TJC121 laboratory at the University of California, Riverside (Choi, Close, Fenton). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see http://www.genome.clemson.edu/projects/barley. To order this clone see http://www.genome.clemson.edu/orders Also see Close T.J., Wing R, Kleinborts A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)" </pre>
BASE COUNT	201 a 251 c 234 g 186 t 3 others
ORIGIN	
Query Match	77.3%; Score 317; DB 10; Length 875;
Best Local Similarity	86.8%; Pred. No. 7e-63;
Matches 347; Conservative 0; Mismatches 53; Indels 0; Gaps 0;	
11	cgcgcatgagggcagaacacttgcctcacaatgacagagaaggggtcccatctctgt 70
143	CGCGGTCTATGGGGCAGAACCTTGGCCCTCAATGACAGAAAGGCTTCCCATCTCTGT 202
71	gtacaacaggaacaacttcaaggtgagagagacgtgagcgttcacaagagagaagaaa 130
203	CTACACAGGACGATCTCAATCAAGGTGAGGAACCGCTCCAGGCTGCCACGACTAGAAAGGAAA 262
131	ccctccgctctacaggtctcattgaccccgctctcttcttgtaagtcattcagaagccagc 190
263	ccctccgctctacaggtctcattgaccccgctctcttcttgtaagtcattcagaagccagc 322

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 03:38:46 ; Search time 6582.52 Seconds

(without alignments)
469,547 Million cell updates/sec

Title: US-09-300-482-27

Perfect score: 229
1 cagaccattttctcgtc.....ctcaattatgggagacaa 229

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pin:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152.2	66.5	566	10	B1972100 sag87c05.
2	139.4	60.9	561	9	BE022117 sm67f12.Y
3	113.4	49.5	532	9	AW719787 L1NEST10A
4	110.4	48.2	430	9	AV406682 AY406682
5	108.6	47.4	630	9	AW695821
6	105	45.9	355	10	BG447744 NF093E04E
7	101	44.1	499	9	BE020243
8	100	43.7	435	10	BE640081
9	96.2	42.0	573	9	AW586409
10	96	41.9	367	10	BF632565
11	96	41.9	430	9	AL368071
12	94	41.0	492	10	B1262817
13	89.6	39.1	420	9	AL382847
14	88.2	38.5	627	10	BF645919
15	87.4	38.2	337	9	AL367204
16	87.4	38.2	344	10	BF636762
17	87.4	38.2	468	10	B1266601

18	87.4	38.2	479	9	AL365745	AL365745 MEB402B10
19	87.4	38.2	481	9	BF639372	BF639372 NF011C10I
20	87.4	38.2	485	10	BF639372	BF639372 NF099D04I
21	87.4	38.2	489	10	B1267224	B1267224 NF116E04P
22	87.4	38.2	533	10	B1264810	B1264810 NF030C11I
23	87.4	38.2	574	10	BF639271	BF639271 NF009P08S
24	87.4	38.2	630	9	AW688619	AW688619 NF066E08I
25	87.4	38.2	646	10	BF642248	BF642248 NF03F09I
26	87.4	38.2	664	10	B1267384	B1267384 EST53038I
27	87.4	38.2	677	10	B1308974	B1308974 EST53038I
28	87.4	38.2	684	10	B1266145	B1266145 NF100H06I
29	85.6	37.4	321	9	AV410206	AV410206 AV410206
30	85.6	37.4	379	9	AV410953	AV410953 AV410953
31	85.6	37.4	394	9	AV425788	AV425788 AV425788
32	85.6	37.4	548	10	B1419927	B1419927 L1NEST42A
33	85.6	37.4	548	10	B1420937	B1420937 L1NEST63C
34	82.2	35.9	606	9	AW299197	AW299197 EST306007
35	81	35.4	559	10	BM093703	BM093703 sa113b10.
36	81	35.4	614	10	BF066944	BF066944 sc35d10.Y
37	80.6	35.2	548	9	AL368600	AL368600 MEB435E04
38	80.4	35.1	670	10	BG598747	BG598747 EST503647
39	79.6	34.8	520	10	B1920404	B1920404 EST540339
40	79.4	34.7	733	9	AI055038	AI055038 coa00002N
41	79	34.5	379	10	BG239631	BG239631 sab72d08.
42	78.4	34.2	151	10	BF645629	BF645629 NF030F01E
43	78	34.1	332	9	BE204727	BE204727 EST397403
44	78	34.1	504	10	BE450814	BE450814 EST401701
45	78	34.1	556	10	B1179863	B1179863 EST520808

ALIGNMENTS

RESULT 1
LOCUS B1972100
DEFINITION B1972100 566 bp mRNA linear EST 29-NOV-2001
ID: Gm-cl084-1257 5' similar to TR:022111 022111 6-PHOSPHOGULCONATE
DEHYDROGENASE ; mRNA sequence.

ACCESSION B1972100
VERSION B1972100
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 566)
AUTHORS Shoemaker,R., Kelm,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Rheising,B., Allen,M., Bowers,
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,
R., Rittler,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,
R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project

TITLE JOURNAL
COMMENT
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available through: ResGen, Invitrogen Corp, 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: c@resgen.com
High quality sequence stop: 421.
Location/Qualifiers
1..566
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl084-1257"


```

/clonelib="Gm-cl084"
/tissue_type="Etiolated hypocotyls (Williams 82)"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed by M. Bhattacharya
from mRNA isolated from etiolated hypocotyls from the
cultivar Williams 82. Tissue was inoculated with
Phytophthora soyae race 1 and tissues were harvested 2 and
4 hours following infection. The library is the pool of
these two time points. Complementary DNA was synthesized
from mRNA using a primer consisting of a poly(dt) sequence
with a XhoI restriction site. EcoRI adapters were ligated
to the blunt-ended cDNA fragments followed by XhoI
digestion. The cDNA fragments were directionally cloned
into the EcoRI-XhoI restriction site of the pBluescript
vector. The ligated cDNA fragments were transformed into
DH10B host cells (Gibco BRL). This library was constructed
by M. Bhattacharya in the laboratory of Dr. Randy
Shoemaker at Iowa State University."

```

```

BASE COUNT      158 a      107 c      140 g      161 t
ORIGIN
Query Match      66.5%; Score 152.2; DB 10; Length 566;
Best Local Similarity 84.7%; Pred. No. 2e-34;
Matches 194; Conservative 0; Mismatches 33; Indels 2; Gaps 2;

```

```

QY 1 cagacctattttcttcatttcctcaatttcagagattatattatgcgtcaacc 60
    |||||||
Db 3 CAGACCTTATTTTCTCTCATTTGATTCAAATTCAGAGATTATATG-GCTCAACC 61
QY 61 cacaacaagaataggcctctgcgtgattgctgttaatggcaaaatctgcgaactcaat 120
    |||||||
Db 62 CACACACAGAAATAGGCTTCTGCTGATTTGGCTGTTATGGCCAAATCTGGCACATCAATAT 121
QY 121 tgcctgaaagggtctcccaattccggttaacaacggaaccttccaaagttattgggc 180
    |||||||
Db 122 TGC-TGAGAAAGCGTTTCCCATTTCTGTTTACACCGAACCACTTCCAGGTTTATGAGA 180
QY 181 cataagacgaagcaaacgaagaagaaccttcaattatggggaaca 229
    |||||||
Db 181 CAGTAGAACGACGCAACAGAGAAATCTTCAGTTTATGCTACCA 229

```

```

RESULT 2
BE022117      561 bp      mRNA      linear      EST 03-DEC-2001
LOCUS      sm67f12.y1 Gm-cl028 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION      Gm-cl028-9000 5' similar to TR:022111 O22111 6-PHOSPHOGLUCONATE
ACCESSION      BE022117
VERSION      BE022117.1 GI:8284549
KEYWORDS      EST.
SOURCE      soybean.
ORGANISM      Glycine max
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseolae;
                Glycine.
REFERENCE      1 (bases 1 to 561)
AUTHORS      Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna
                'A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
                Wylie,T., Underwood,R., Steptoe,M., Theising,B., Allen,M., Bowers,
                'Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
                'R., Riter,E., Kohn,S., Shin,T., Jackson,T., Cardenas,M., McCann
                'R., Waterston,R. and Wilson,R.
                Public Soybean EST Project
                Unpublished (1999)
                Contact: Shoemaker R/Public Soybean EST Project
                Public Soybean EST Project
                Washington University School of Medicine
                4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                Tel: 314 286 1800

```

```

FEATURES
    source
        Fax: 314 286 1810
        Email: est@watson.wustl.edu
        This clone is available through: Reagen, Invitrogen Corp. 2130
        South Memorial Parkway Huntsville, AL 35801 for further information
        call: (800)-533-4363 or contact via email: cu@reagen.com
        Insert length: 1893      Std Error: 0.00
        High quality sequence stop: 398.
        Location/Qualifiers
            1..561
                /organism="Glycine max"
                /db_xref="taxon:3847"
                /clone="GENOME SYSTEMS CLONE ID: Gm-cl028-9000"
                /clone_lib="Gm-cl028"
                /tissue_type="roots of 'Superpod' plants"
                /lab_host="DH10B"
                /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                XhoI; The mRNA was isolated from roots of Glycine max
                'Superpod' plants generously donated by Dr. Gary Stacey.
                The seedlings were inoculated with Bradyrhizobium
                japonicus, strain USDA110 prior to harvest. Stragene's
                cDNA synthesis kit (catalog number 200401) was used to
                synthesize the cDNA. First-strand synthesis was performed
                with 5-methyl dCTP, hence the ligated cDNA was
                hemimethylated. A modification of Stragene's
                first-strand synthesis primer was used. An 'anchor'
                nucleotide (V-A-C, or G) was added to the 3' end of the
                primer (GAGAGAGAGACAGAGAGACATGTCGAG(1)18V) to anchor
                the primer at the 5' end of the poly(A) tract. After
                second-strand synthesis, the cDNA ends were filled in with
                cloned Pfu DNA polymerase, ligated to EcoRI adapters and
                subsequently phosphorylated. The XhoI site within the
                first-strand synthesis primer was then restricted by
                digestion with XhoI; all XhoI sites in the cDNA would be
                protected by their hemimethylated status. The cDNA
                constructs were size fractionated with a 500bp cutoff,
                using GIBCOBRL Life Technologies' cDNA size fractionation
                column. The column eluent was then ligated into
                Stragene's pBluescript II SK(+) that has been digested with EcoRI
                (pBluescript II SK(+)) that has been digested with EcoRI
                and XhoI, and phosphorylated by Stragene). Both the
                white and blue colonies appear to contain recombinant
                plasmids with cDNA inserts, based on size (n=25). This
                library was constructed by Dr. Paul Keim and Dr. Virginia
                Corryell."

```

```

BASE COUNT      158 a      115 c      130 g      158 t
ORIGIN
Query Match      60.9%; Score 139.4; DB 9; Length 561;
Best Local Similarity 81.2%; Pred. No. 1.2e-30;
Matches 186; Conservative 0; Mismatches 41; Indels 2; Gaps 2;

```

```

QY 1 cagacctattttcttcatttcctcaatttcagagattatattatgcgtcaacc 60
    |||||||
Db 7 CAGACCTTATTTTCTCTCATTTGCTTCAATTCAGAGATTATATG-GCTCAACC 65
QY 61 cacaacaagaataggcctctgcgtgattgctgttaatggcaaaatctgcgaactcaat 120
    |||||||
Db 66 CTCACACAGAAATAGGCTTCTGCTGACCTGGCTGTTATGGCCAAATCTAGCACATCAATAT 125
QY 121 tgcctgaaagggtctcccaattccggttaacaacggaaccttccaaagttattgggc 180
    |||||||
Db 126 TGC-TGAGAAAGCGTTTCCCATTTCTGTTTATACCGAACACTTCCAAAGGTGATGAGA 184
QY 181 cataagacgaagcaaacgaagaagaaccttcaattatggggaaca 229
    |||||||
Db 185 CTGTTAGAACGACGCAACAGAGAAATCTTCAGTTTATGCTACCA 233

```

```

RESULT 3
AW719787      532 bp      mRNA      linear      EST 19-APR-2000
LOCUS      AW719787
DEFINITION      LJWEST1044r Lotus japonicus nodule library 5 and 7 week-old Lotus

```


ACCESSION	japonicus cDNA 5', mRNA sequence.
VERSION	AJ719787
KEYWORDS	EST.
SOURCE	Lotus japonicus.
ORGANISM	Lotus japonicus
REFERENCE	1 (bases 1 to 532)
AUTHORS	Colebatch,G., Freund,S., Trewaskis,B and Udvardi,M.
TITLE	Lotus japonicus root nodule ESTs: tools for functional genomics
JOURNAL	Unpublished (2000)
COMMENT	Contact: Udvardi MK Molecular Plant Nutrition Max Planck Institute of Molecular Plant Physiology Am Muehlenberg 1, 14476 Golm, Germany Fax: 49 331 567 8250 Email: udvardi@mpln-mpg.mpg.de Seq primer: T7 High quality sequence stop: 532. Location/Qualifiers 1..532 /organism="Lotus japonicus" /cultivar="Gift (B-129)" /db_xref="taxon:34305" /clone_lib="Lotus japonicus nodule library 5 and 7 week-old" /dex_stage="5 and 7 week-old plants" /note="Organ: Nodule; Vector: pSPORT1; Site_1: Salt; Site_2: NaCl; The library was prepared using mRNA extracted from nodules of 5 and 7 week-old Lotus plants. Nodules were induced by, and contained Mesorhizobium strain R7A."
BASE COUNT	151 a 102 c 137 g 140 t 2 others
ORIGIN	
Query Match	49.5%; Score 113.4; DB 9; Length 532;
Best Local Similarity	74.2%; Pred. No. 5.4e-23;
Matches 170; Conservative	0; Mismatches 56; Indels 3; Gaps 2;
OY	3 gacctatttttctgtcatttgctcaaatctcgagatattatgcgtcaacca 62
Db	
	2 GACCTCAATTTTGGTGATTTCGATCATTTCAGAAATAATTAATGCTCAACC61
OY	63 c--acaagaataggcccttcgtgatcgtcgttaataaggcaaaatcgcgcactaatat 120
Db	62 TCCTTAGCGAATAGGCCCTGTGAGCTGCCTGTATGGCCCAATCTTCACACTCAATAT 121
OY	121 tgcttgaagaaggcttcccaattccggttaacaacggaaccattccaagtatttgyc 180
Db	
	122 TGC-TGACAAAAGGCTTCCCATTTCTGTTAACACCGACACCAATCCAGGTGATGAGA 180
OY	181 cataagacgaagcaaaccaaggaagaacccattcaattatgggagaaca 229
Db	
	181 CAGTTGAACGGGCACAACACGAGAAGAACCTTCCAGTGTATGCGTACCA 229
RESULT	4
LOCUS	AV406682 430 bp mRNA linear EST 23-MAY-2000
DEFINITION	japonicus young plants (two-week old) Lotus
ACCESSION	AV406682
VERSION	AV406682.1 GI:7719536
KEYWORDS	EST.
SOURCE	Lotus japonicus.
ORGANISM	Lotus japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;	

REFERENCE	1 (bases 1 to 430)
AUTHORS	Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE	Generation of 7137 non-redundant expressed sequence tags from a legume, Lotus japonicus
JOURNAL	Legume, Lotus japonicus
MEDLINE	Lotus Res. 7 (2), 127-130 (2000)
COMMENT	20277479
CONTACT:	Yasukazu Nakamura
THE FIRST LABORATORY FOR PLANT GENE RESEARCH	
KAZUSA DNA RESEARCH INSTITUTE	
YANA 1532-3, KISARAZU, CHIBA 292-0812, JAPAN	
EMAIL: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.	
FEATURES	
SOURCE	Location/Qualifiers
	1..430
	/organism="Lotus japonicus"
	/db_xref="taxon:34305"
	/clone="MWL008a02.x"
	/clone_1db="Lotus japonicus young plants (two-week old)"
	/clon_stage="young plants (two-week old)"
	/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI; isolate=Miyakojima MG-20"
BASE COUNT	127 a 87 c 102 g 114 t
ORIGIN	
Query Match	48.2%; Score 110.4; DB 9; Length 430;
Best Local Similarity	73.9%; Pred. No. 4.1e-22;
Matches 167; Conservative 0; Mismatches 56; Indels 3; Gaps 2;	
OY	6 ctatttttttctgtcatttgcctcaaatctcaagagatattatgctcaccaca--c 63
Db	1 CTCATTTTTCCTGTTCATTCGATTCATTAATTCGAGAAATAATTAATGCTCAACCAACTCT 60
OY	64 aacagaatagacctgtcgtgattggccgttaatggcaaaatctgcgaccatcatatgc 123
Db	61 TAGCGAATAGAGCCCTGCTGAGCTGACCTGGCTGTTATGGGCGCAAAATCTTGCACTCAATATGCG 120
OY	124 ttgaaaggcttcccaattcccgyltaacaacggaacatttccaaagttattggccat 183
Db	121 -TGAGAAAGGCTTCCCTTCCTGTTTCAACCGGACACCATCCAAAGGTGATGAGACAG 179
OY	184 aagacgaagcaaacaccggaaggaaccttcaattatgagggaacaa 229
Db	180 TTGAACGGGCACAAACGAGAGAAACCTTCACAGTATATGCTACCA 225
RESULT 5	
AM695821	630 bp mRNA linear EST 20-DEC-2000
LOCUS	NP099E09S1F1070 Developing stem Medicago truncatula cDNA clone
DEFINITION	NP099E09S1F1070 Developing stem Medicago truncatula cDNA clone
ACCESSION	AM695821
VERSION	AM695821.2 GI:11932865
KEYWORDS	EST.
SOURCE	barrel medic.
ORGANISM	Medicago truncatula
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
	Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
	Medicago.
REFERENCE	1 (bases 1 to 630)
AUTHORS	He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores,H.R., Imman,J.T., Weller,J.W., May,G.D. and Dixon,R.A.
TITLE	Expressed Sequence Tags from the Samuel Roberts Noble Foundation
JOURNAL	Medicago truncatula stem library
COMMENT	Unpublished (2000)
	On Apr 14, 2000 this sequence version replaced gi:7570583.
	Contact: Dixon RA
	Plant Biology Division
	The Samuel Roberts Noble Foundation
	2510 Sam Noble Parkway, Ardmore, OK 73402, USA
	Tel: 580 221 7302

FEATURES High quality sequence stop: 382.
 Location/Qualifiers

Tel: 580-223-5810
 Fax: 580-221-7380
 Email: mjharison@noble.org
 Noble EST name: N253596
 TIGR sequence name: MDA086TK
 More information is available at:
 'http://chryslr.tamu.edu/medicago/'
 Seq primer: Skmod (CTA GAA GTA gAT CC).
 Location/Qualifiers

FEATURES

source

1..573

/organism="Medicago truncatula/Glomus versiforme mixed EST library"

/cultivar="Medicago truncatula genotype A17"

/db_xref="taxon:119092"

/clone="PMHAM-5403"

/clone.lib="MHAM"

/tissue_type="roots colonized with Glomus versiforme"

/dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days post-inoculation with Glomus versiforme. The library was made from a mixture of RNA from each of these stages."

/lab_host="E. coli strain XLOLR"

/note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2: XhoI; cDNA was prepared from polyA+ enriched RNA from roots harvested at 10, 17, 22, 31 and 38 days post-inoculation with Glomus versiforme. The cDNA was directionally ligated into the UniZap XR vector from StrataGene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XLOLR cells."

BASE COUNT 164 a 115 c 139 g 154 t 1 others

ORIGIN

Query Match 42.0%; Score 96.2; DB 9; Length 573;

Best Local Similarity 74.6%; Pred. No. 6.4e-18;

Matches 147; Conservative 0; Mismatches 48; Indels 2; Gaps 2;

QY 21 attgctcaatttcaggagatcaattatgctgctcaaccacaagaatagccttg 80

Db 7 AATCTCTTCAATTTTCAGGAATACATTATG-GCTCAACCACTTACAGATAGCGCTTG 65

QY 81 ctgattgctgttaattggaacatctgacccaatatgcttgtaaaaggcttcga 140

Db 66 CCGGACTGGCTGTATGCGCAAAATCTCGACATCATATTGC-GGACAAAGATTCCCA 124

QY 141 attcggtttaacacgaagacattccaaggtatctggccataagaagcaaacag 200

Db 125 ATTCTGTATTATACAGAAACATCAAGGTGACGAACCTGTGAAACGACAAACAG 184

QY 201 gaaggaacctcaatt 217

Db 185 GAGGGAATCTTCCACT 201

RESULT 10

Bf632565

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

367 bp mRNA linear EST 19-DEC-2000

NF032C07/DTJF1051 Drought Medicago truncatula cDNA clone NF032C07/DT

5', mRNA sequence.

Bf632565

EST.

barrel medic.

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;

Medicago.

1 (bases 1 to 367)

Torrez-Varez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,

Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

JOURNAL

Medicago truncatula drought library
 Unpublished (2000)

COMMENT

Contact: May GD

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7391

Fax: 580 221 7380

Email: gdmay@noble.org

Insert Length: 367

Std Error: 0.00

Plate: 032

row: C

column: 07

Seq primer: TCACACAGGAACGCTATGAC.

FEATURES

source

1..367

/organism="Medicago truncatula"

/db_xref="taxon:3880"

/clone="NF032C07DT"

/clone.lib="Drought"

/tissue_type="Plantlets"

/dev_stage="Pooled timepoints"

/note="Vector: Lambda Zap; Contains a mixture of entire plantlets harvested in a series of days-post-watering timepoints."

BASE COUNT 117 a 75 c 80 g 95 t

ORIGIN

Query Match 41.9%; Score 96; DB 10; Length 367;

Best Local Similarity 75.5%; Pred. No. 7.1e-18;

Matches 145; Conservative 0; Mismatches 45; Indels 2; Gaps 2;

QY 26 ctccaatttcaggagatcaattatgctgctcaaccacaagaatagccttgctga 85

Db 1 CTTCATTTTCAGGAATACATTATG-GCTCAACCACTTACAGATAGCGCTTGCCGGA 59

QY 86 ttgctgttaattggaacatctgacccaatatgcttgtaaaaggcttcgaatcc 145

Db 60 CTGGCTGTATGCGCCAAATCTCGACATCAATTATGC-CGACAAAGATTCCCAATTTC 118

QY 146 ggttaacacggaacattccaaggttctggccataagaagcaaacaggaagg 205

Db 119 TGTATTACAGAAACATCAAGGTGACGAACCTGTGAAACGACAAACAGAGGG 178

QY 206 aaacctcaatt 217

Db 179 AATCTTCCACT 190

RESULT 11

AL368071

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

430 bp mRNA linear EST 03-AUG-2000

MTBA22A07F1 MTBA Medicago truncatula cDNA clone MTBA22A07 T3, mRNA

sequence.

AL368071

EST.

barrel medic.

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;

Medicago.

1 (bases 1 to 430)

Journet,E.P., Crespeau,H., van-Tulmen,D., Gouzy,J., Jallion,O.,

Niebel,A., Carreau,V., Chataigner,O., Kahn,D., Gianinazzi-Pearson,

V. and Gamas,P.

Medicago truncatula ESTs from nitrogen-starved roots

Unpublished (2000)

Genoscope Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Contact : Pascal Gamas and Etienne-Pascal Journet, Laboratoire de

Biologie Moléculaire des Relations Plantes-Microorganismes,
CNRS-INRA, BP 27 31326 Castanet-Tolosan Cedex, France (Email :
Mt-est@toulouse.inra.fr Website :
http://sequence.toulouse.inra.fr/Mtruncatula.html).

FEATURES
source
1. .430
Location/Qualifiers

/organism="Medicago truncatula"
/cultivar="Jemalong"
/db_xref="taxon:3880"
/clone="MCBA22A07"
/clone_lib="MTBA"
/tissue_type="root tips"
/dev_stage="harvested after 3 days of N-starvation"
/note="Vector: pBluescript psk; Site_1: EcoRI; Site_2:
XhoI; Plants were grown in an aeroponic chamber for 14
days on nitrogen-rich medium followed by 3 days on N-free
medium. RNA was extracted from root tips (1-3 cm). cDNA
was prepared from polyA+ enriched RNA. The cDNA was
directionally ligated into Uni-ZapR vector from
Stratagene and packaged using Gigapack Gold packaging
extracts. Plasmids containing cDNA inserts were
mass-excised from phage stocks using ExAssit helper phage
and propagated in SOLR cells. Clone ordering and
sequencing was performed by the Centre National de
Séquence (Genoscope, Evry, France)."

BASE COUNT
ORIGIN
131 a 85 c 103 g 111 t

Query Match
Best Local Similarity 75.5%; Pred. No. 7.2e-18; Length 430;

Matches 145; Conservative 0; Mismatches 45; Indels 2; Gaps 2;

QY 26 ctccaattcagaagatataatgctcgaaccacacagaagaatagccttcgga 85
||||| ||||||| || ||||||| ||||||| ||||||| ||||||| ||
Db 2 CTTCATTTCAGAGAATACATTATG-GCTCAACCACTTACAGAATAGCGCTTGCGGA 60
QY 86 ttgctgttaattggcgaactctgcacataatctgctgaagggtcccaattcc 145
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 61 CTGCGCTTTATGGGCCAAATCTGCACATCAATATTCG-CGACAAAGATTCCTCAATTTTC 119
QY 146 ggttaacacgaagacattcccaagtatttgccataagaagacaaacgaagg 205
||||| ||||||| || ||||||| || || ||||||| ||||||| ||
Db 120 TGTATTAAACAGAACACATCAAGTTGACGAACCTGTGAACGACGACGAGG 179
QY 206 aaactcatt 217
||| ||||| ||
Db 180 AAATCTTCCACT 191

RESULT 12

LOCUS BI262817 492 bp mRNA linear EST 18-JUL-2001
DEFINITION NF091D06EC1F1058 Elicited cell culture Medicago truncatula cDNA
clone NF091D06EC 5', mRNA sequence.

ACCESSION BI262817
VERSION BI262817.1 GI:14863469
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula

REFERENCE 1 (bases 1 to 492)
AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
C.U., Imman,U.T., May,G.D. and Dixon,R.A.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
JOURNAL Medicago truncatula elicited cell culture library
COMMENT Unpublished (2001)
Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 492 Std Error: 0.00
Plate: 091 row: D column: 06
Seq primer: TCACACAGGAACAGCTATGAC.
Location/Qualifiers

FEATURES
source
1. .492
Location/Qualifiers

/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF091D06EC"
/clone_lib="Elicited cell culture"
/tissue_type="Cell cultures derived from root tissues"
/dev_stage="Cell suspensions were subcultured every 14
days. Cells were induced six days after subculture"
/note="Vector: Lambda Zap; Cells were induced with yeast
cell wall extracts equivalent to 50ug/ml glucose in the
final concentration. Samples were taken at 0.5, 1, 12 and
24 hours after induction. Equal amounts of RNA from each
time point were pooled and used for mRNA isolation."

BASE COUNT
ORIGIN
143 a 98 c 119 g 128 t 4 others

Query Match
Best Local Similarity 74.5%; Pred. No. 2.8e-17; Length 492;

Matches 143; Conservative 0; Mismatches 47; Indels 2; Gaps 2;

QY 26 ctccaattcagaagatataatgctcgaaccacacagaagaatagccttcgga 85
||||| ||||||| || ||||||| ||||||| ||||||| ||||||| ||
Db 5 CTTCATTTCAGAGAATACATTATG-GCTCAACCACTTACAGAATAGCGCTTGCGGA 63
QY 86 ttgctgttaattggcgaactctgcacataatctgctgaagggtcccaattcc 145
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 64 CTGCGCTTTATGGGCCAAATCTGCACATCAATATTCG-CGACAAAGATTCCTCAATTTTC 122
QY 146 ggttaacacgaagacattcccaagtatttgccataagaagacaaacgaagg 205
||||| ||||||| || ||||||| || || ||||||| ||||||| ||
Db 123 TGTATTAAACAGAACACATCAAGTTGACGAACCTGTGAACGACGACGAGG 182
QY 206 aaactcatt 217
||| ||||| ||
Db 183 AAATCTTCCACT 194

RESULT 13

LOCUS AL382847 420 bp mRNA linear EST 03-AUG-2000
DEFINITION MBBC10C10F1 MtBC Medicago truncatula cDNA clone MtBC10C10 T3, mRNA
sequence.

ACCESSION AL382847
VERSION AL382847.1 GI:9682598
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula

REFERENCE 1 (bases 1 to 420)
AUTHORS Journe!t,E.P., Crespeau,H., van-Tulnen,D., Gouzy,J., Jallou,O.,
Njebel,A., Carreau,V., Chataigner,O., Kahn,D., Glaninaazi-Pearson
V. and Gamas,P.
TITLE Medicago truncatula ESTs from endomycorrhizal roots
JOURNAL Unpublished (2000)
CONTACT: Genoscope
Genoscope - Centre National de Séquencage
BP 191 91006 EVRY cedex - France
Email: sequef@genoscope.cns.fr. Web : www.genoscope.cns.fr
Contact : Pascal Gamas and Etienne-Pascal Journe!t, Laboratoire de
Biologie Moléculaire des Relations Plantes-Microorganismes,
CNRS-INRA, BP 27 31326 Castanet-Tolosan Cedex, France (Email :

FEATURES
source

Location/Qualifiers

http://sequence.toulouse.inra.fr/Mtruncatula.html

1..420

/organism="Medicago truncatula"

/cultivar="Jemalong"

/db_xref="taxon:3880"

/clone="MTBC10C10"

/clone.lib="MTBC"

/tissue_type="arbuscular mycorrhiza"

/dev_stage="harvested 3 weeks post inoculation with Glomus intraradices"

/note="Vector: pBluescript PSK; Site 1: EcoRI; Site 2: XhoI; M. truncatula sterilised seeds were germinated for 72h at 25 C, before transplanting into a 1/3 Epioses soil

: 2/3 calcined Terragreen mix in the presence of onion root fragments colonized by the arbuscular mycorrhizal fungus Glomus intraradices (Schenck & Smith, isolate LP8). The plants were watered every day and twice a week with a modified nutrient long Ashton solution without phosphate but with a high level of nitrate. After 3 weeks RNA was extracted from whole root systems. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into Uni-zap XR vector from Stratagene and packaged using GigaPack Gold packaging extracts. Plasmids containing cDNA inserts were mass-excised from phage stocks using Exasit helper phage and propagated in SOLR cells. Clone ordering and sequencing was performed by the Centre National de Sequencage (Genoscope, Evry, France). Note : EST may be of fungal origin."

BASE COUNT 128 a 83 c 102 g 106 t 1 others

ORIGIN

Query Match 39.1%; Score 89.6; DB 9; Length 420;
Best Local Similarity 74.6%; Pred. No. 5.6e-16;
Matches 138; Conservative 0; Mismatches 45; Indels 2; Gaps 2;

Qy 33 ttctcaggagatgaatgaatgctgaacccacaagaataagccttgctgagatgctg 92
|||||
Db 1 TTTTCAGGAATACATATAG-GCTACACCACTTACAGAAATAGGCGCTCGGACTGCTG 59
Qy 93 ttaatgggcaaatctggcactcaatatgtctgaagaagcctcccaatccggttaac 152
|||||
Db 60 TTATGGGCGCAAAATCTCGCACTCAATATGCGACNAAGATTTCCCAATTTCTGTTAT 118
Qy 153 aaggaaccatttccaaagtattatgggcataagaagcaagaacccaggaagaaacct 212
|||||
Db 119 AACGAAACAAACATCAAAAGTTGACGAAACCTGTGAACGACCAAAACAGAGGAAATCTT 178
Qy 213 caatt 217
|||
Db 179 CCACT 183

RESULT 14 627 bp mRNA linear EST 20-DEC-2000
BF645919
LOCUS NF040E03ECJF1022 Elicited cell culture Medicago truncatula cDNA
DEFINITION clone NF040E03EC 5', mRNA sequence.
ACCESSION BF645919
VERSION BF645919.1 GI:11911048
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 627)
Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation -

JOURNAL
COMMENT Center for Medicago Genomics Research
Unpublished (2000)
Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 627 Std Error: 0.00
Plate: 040 row: E column: 03
Seq primer: TCACACGGAACACGCTATGAC.

FEATURES
source

Location/Qualifiers

1..627

/organism="Medicago truncatula"

/db_xref="taxon:3880"

/clone="NF040E03EC"

/clone.lib="Elicited cell culture"

/tissue_type="Cell cultures derived from root tissues"

/dev_stage="Cell suspensions were subcultured every 14 days. Cells were induced six days after subculture"

/note="Vector: Lambda Zap; Cells were induced with yeast cell wall extracts equivalent to 50ug/ml glucose in the final concentration. Samples were taken at 0.5, 1, 12 and 24 hours after induction. Equal amounts of RNA from each time point were pooled and used for mRNA isolation."

BASE COUNT 180 a 123 c 148 g 174 t 2 others

ORIGIN

Query Match 38.5%; Score 88.2; DB 10; Length 627;
Best Local Similarity 73.2%; Pred. No. 1.5e-15;
Matches 153; Conservative 0; Mismatches 53; Indels 3; Gaps 3;

Qy 11 ttcttctgcatcttgcctcaaatctcagaagatgaatgctgctcaacc-cacacaag 69
|||||
Db 3 TTTTTCACCTAATCTCTCAATTTTCAGAAATACATATGCTGCTCAACCACTTACAG 62
Qy 70 aatagcctctgctgattgg-cgttaetgggcaaatctggcactcaatatgtctgaa 128
|||||
Db 63 AATAGGCTTGGCGGAGCTGCTGTTATGCGCCAAATCTCGCACTCAATATGCG-CGAC 121
Qy 129 aaggcttcccaatctcggttaacaacgaacatttccaaagtattatgggcataaag 188
|||||
Db 122 AAGGATTTCCCAATTTCTGTTATAACAGAACACATCAAAAGCTTGACGAAACGTGGA 181
Qy 189 gaagcaaccaggaagaaaccttcaatt 217
|||||
Db 182 CGAGCAAAACAGAGGAGGAATCTTCACCT 210

RESULT 15 337 bp mRNA linear EST 03-AUG-2000
AL367204
LOCUS MTBA12H03F1 MTBA Medicago truncatula cDNA clone MTBA12H03 T3, mRNA
DEFINITION sequence.
ACCESSION AL367204
VERSION AL367204.1 GI:9666957
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 337)
Journet,E.P., Crespeau,H., van-Tuluen,D., Gouzy,J., Jaillon,O.,
Nebel,A., Carreau,V., Chabagnier,O., Kahn,D., Gianinazzi-Pearson
,V. and Gamas,P.
Medicago truncatula ESTs from nitrogen-starved roots
Unpublished (2000)
Contact: Genoscope
Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Contact : Pascal Gamas and Etienne-Pascal Journet, Laboratoire de
 Biologie Moléculaire des Relations Plantes-Microorganismes,
 CNRS-INRA, BP 27 31326 Castanet-Tolosan Cedex, France (Email :
 Mt-est@toulouse.inra.fr Website :
 http://sequence.toulouse.inra.fr/Mtruncatula.html).

FEATURES

source

```
1. .337
  /organism="Medicago truncatula"
  /cultivar="Jemalong"
  /db_xref="taxon:3880"
  /clone="MtBA12H03"
  /clone_11b="MtBA"
  /tissue_type="root tips"
  /dev_stage="harvested after 3 days of N-starvation"
  /note="Vector: pBluescript PSK; Site.1: EcoRI; Site.2:
  XhoI; Plants were grown in an aeroponic chamber for 14
  days on nitrogen-rich medium followed by 3 days on N-free
  medium. RNA was extracted from root tips (1-3 cm). cDNA
  was prepared from polyA+ enriched RNA. The cDNA was
  directionally ligated into Uni-zapXR vector from
  Stratagene and packaged using Gigapack Gold packaging
  extracts. Plasmids containing cDNA inserts were
  mass-excised from phage stocks using ExsacIII helper phage
  and propagated in SOLR cells. Clone ordering and
  sequencing was performed by the Centre National de
  Sequencage (Genoscope, Evry, France)."
```

BASE COUNT 108 a 71 c 72 g 86 t

ORIGIN

Query Match

38.2%; Score 87.4; DB 9; Length 337;

Best Local Similarity 74.5%; Pred. No. 2.4e-15;

Matches 123; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

```
QY 53 gctcaaccacaagaatagcctgtgctgtgctgtaattgaggaactctggca 112
   ||||||| ||||||| ||||||| ||||||| ||||||| ||
DB 16 GCTCAACCACTTACAGAAATAGCGCTTGGCTGTTATGGCCAAATCTCGCA 75
   ||||||| ||||||| ||||||| ||||||| ||||||| ||
QY 113 ctaatatgtctgaaaggctcccaatcgggttaacaacggaaccattccaagt 172
   ||||||| || ||||||| || ||||||| || ||||||| || |||||||
DB 76 CTCATATTTTC-CGACCAAGGATTCCCAATTTCTGTTTATACAGACATCAAGGT 134
   ||||||| || ||||||| || ||||||| || ||||||| || |||||||
QY 173 tattggccaataagacgaacgaacgaaggaaaccttcaatt 217
   | | | | | | | | | | | | | | | | | | | | | |
DB 135 TGACGAAACTGTGACAGACGACAAACGAGGAAATCTTCACT 179
   || || || || || || || || || || || || || || || ||
```

Search completed: July 3, 2002, 03:38:50
 Job time: 43310 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2002, 15:37:00 ; Search time 5254.53 Seconds

(without alignments)
927.939 Million cell updates/sec

Title: us-09-300-482-1

Perfect score: 233

Sequence: 1 gttttgcagcttagtagaat.....ttgacaagggtattgttg 233

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :
1: GenEmbl:*
2: gb_hhg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: gb_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hhg_hum:*
31: em_hhg_inv:*
32: em_hhg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	131.6	56.5	1904	AB029456
2	131.6	56.5	1924	AB029455
3	131.6	56.5	1956	AB029454
4	106.8	45.8	59348	H0811E11
5	106.8	45.8	121690	OSJN00125
6	106.8	45.8	151889	OSJN00053
7	106.6	45.8	1785	MSU18238
8	101.8	43.3	1689	STG6PDH
9	100.8	43.3	1957	AF012863
10	97	41.6	1821	NTTCG9
11	96	41.2	1848	AF012862
12	95.4	40.9	1810	AY065054
13	95.4	40.9	1852	NTTCG6
14	95.4	40.9	1862	ATH010970
15	93.8	40.3	1690	ATH010971
16	78	33.5	63604	AP000381
17	77.8	33.4	1777	AF097663
18	74.6	32.0	14323	AB015470
19	39	16.7	37184	HS09D4
20	39	16.7	180008	AC022962
21	38.4	16.5	168106	AC099540
22	36.8	15.8	110000	LMFCHR34_11
23	36.8	15.8	110000	LMFCHR34_12
24	35.8	15.4	117840	AL162400
25	35.8	15.4	147990	AC026144
26	35.4	15.2	63441	AL356502
27	35	15.0	166239	AC078965
28	35	15.0	168822	AC007511
29	34.8	14.9	164012	AC009386
30	34.8	14.9	187258	AC018976
31	34.6	14.8	597	AX141463
32	34.6	14.8	2945	AF269329
33	34.6	14.8	2945	AX144649
34	34.6	14.8	3825	AF269754
35	34.6	14.8	3825	AX145072
36	34.6	14.8	182727	AC068066
37	34.4	14.8	10286	AX277850
38	34.4	14.8	10286	AX323533
39	34.4	14.8	174132	AC095090
40	34.4	14.8	174344	AC094634
41	34.2	14.7	150876	AC108761
42	34.2	14.7	163166	AL356632
43	34	14.6	137480	OSJN00219
44	34	14.6	180844	AL606536
45	33.8	14.5	93830	AC097789

ALIGNMENTS

RESULT	1	1904 bp	mrna	linear	PLN 06-OCT-2000
AB029456	AB029456	Triticum aestivum g6pdh mRNA for glucose-6-phosphate dehydrogenase, complete cds, clone:Ta9p03.			
LOCUS	AB029456				
DEFINITION	AB029456				
ACCESSION	AB029456				
VERSION	AB029456.1				
KEYWORDS	glucose-6-phosphate dehydrogenase.				
SOURCE	Triticum aestivum (cultivar:Chinese spring, isolate:root) seedling				
ORGANISM	Triticum aestivum				
REFERENCE					
AUTHORS	Nemoto,Y., Kawakami,N. and Sasakuma,T.				
TITLE	Isolation of novel early salt-respondering genes from wheat (Triticum aestivum L.) by differential display				
JOURNAL	Theor. Appl. Genet. 98, 673-678 (1999)				
REFERENCE	2 (sites)				
AUTHORS	Nemoto,Y. and Sasakuma,T.				
TITLE	Specific expression of glucose-6-phosphate dehydrogenase (G6PDH) gene by salt stress in wheat (Triticum aestivum L.)				

JOURNAL Plant science (Shannon, Ireland) 158 (1-2), 53-60 (2000)
 PUBMED 10996244
 REFERENCE 3 (basses 1 to 1904)
 AUTHORS Nemoto, Y
 TITLE Direct Submission
 JOURNAL Submitted (01-JUL-1999) yasue Nemoto, Yokohama City University,
 244-0013, Japan (E-mail:nemoto@yokohama-cu.ac.jp,
 tel:81-45-820-1902, fax:81-45-820-1901)

BASE COUNT	ORIGIN
543 a	414 c
464 g	483 t

Query Match	56.58;	Score 131.6;	DB 8;	Length 1904;
Best Local Similarity	76.28;			

Matches	100%	Conservative	0	Mismatches	49	Indels	6	Gaps	1
QY	2	tttttcagcatgtagaataatgttttagtgcctctcatatagatgaggttgaaagattgagta	61						
Db	447	TTCTTTCGCAATTATATAAATAATGCTCACTGGTTCCTATGACAGTGGAAAGATTTCGAAAA	506						
QY	62	ttgaatlgaggaacatctctgtagatgagacttcagaaacaatgactcgggaagctccgcg	121						
Db	507	CTGAAACAAGGAANAATATCAGATTATGAGANGTCA-----ACAACCTCAGGAAGCTTCCTCT	560						
QY	122	agattattttatttggcattgcctcccatcagtcctacccatcagatctgcagatgtaaa	181						
Db	561	AGGGCTCTTTAATTTGGCATTGGCTCCATCTGTCCTACCCCTTCAGTGGCAAAATGATGCCA	620						
QY	182	tcaatactgcatgagccatctctcacacacccggtttgacaaggttaatgctg	233						
Db	621	ACATATTGCATGAGTCCCAACTTCTTCGCCCTGGATGGAGCTACGATAGTAAATGTTTG	672						

RESULT	2
AB029455	
LOCUS	1924 bp mRNA linear PLN 06-OCT-2000
DEFINITION	Triticum aestivum g6pdh mRNA for glucose-6-phosphate dehydrogenase.
ACCESSION	complete cds, clone:Tgapd2.
VERSION	AB029455
KEYWORDS	AB029455.1 GI:8918503
SOURCE	glucose-6-phosphate dehydrogenase. Triticum aestivum (cultivar:Chinese spring; isolate:root) seedling cDNA to mRNA, clone:Tgapd2. Triticum aestivum
• ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE AUTHORS TITLE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum. 1 (sites)
JOURNAL REFERENCE AUTHORS TITLE	Nemoto, Y., Kawakami, N. and Sasakuma, T. Isolation of novel early salt-responsive genes from wheat (<i>Triticum aestivum</i> L.) by differential display Theor. Appl. Genet. 98, 673-678 (1999) 2 (sites)
JOURNAL PUBLISHED REFERENCE AUTHORS TITLE	Nemoto, Y. and Sasakuma, T. Specific expression of glucose-6-phosphate dehydrogenase (G6PDH) gene by salt stress in wheat (<i>Triticum aestivum</i> L.) Plant science (Shannon, Ireland) 158 (1-2), 53-60 (2000) 10996244 3 (bases 1 to 1924)
JOURNAL PUBLISHED REFERENCE AUTHORS TITLE	Nemoto, Y. Direct Submission Submitted (01-JUL-1999) Yasuo Nemoto, Yokohama City University, Kihara Institute for Biological Research, Maika, Totsuka, Yokohama 244-0813, Japan (E-mail:nemoto@yokohama-cu.ac.jp, Tel:-81-45-820-1902, Fax:81-45-820-1901) Location/Qualifiers
FEATURES	

BASE COUNT	545 a	417 c	475 g	487 t
ORIGIN				

Query Match	56.5%	Score 131.6;	DB 8;	Length 1924;
Best Local Similarity	76.3%	Pred. No. 3.1e-29;		
Matches 177; Conservative	0;	Mismatches 49;	Indels 6;	Gaps 11

QY	2	tttttgcagcttagtagaataatgttagtgcctccctatgataagagtgtaaggaatttcagta	61
Db	455	ttctttgcgaattgatgataaattatgtcagtgctgcttccctatgatcagtgagaaaggtttgaaaa	514
QY	62	ttagatggaggaacatctcagatgatgagctctcagaaaaaatgctctgggaagctaccgc	121
Db	515	ctgaacaaaggaattatcagatttatgagatgca-----aaccaactaggaaggtccct	568
QY	122	agatattattatcttgacatcgcctccalcagctccaccatcagatctcgagatgataga	181
Db	569	aggcctcttttatttggcattgcctccatcctatgcttacccttcacgtgccaataatgatccga	628
QY	182	tcatattgcagatccatcttcaacacacggcttggaagaaggtattgttg	233
Db	629	acattattgcacagatcccaactctccggctcgatgagcactagatattattgttg	680

RESULT 3

	gene	23285..29525 /gene="H081IE11.4" join(23285..23479,29436..29525) /gene="H081IE11.4"
CDS		/product="hypothetical protein" /protein_id="CAC09488.1" /db_xref="GI:10241650" /translation="MEAEHPHQITVTTTTSTSLCPRRRKGDDDEAAHLVFPMDL DSEAAAAAHHQOQQOOQTTLVGCTATCIVLPENGEARTRHQVNNNYH" 23480..29435 /gene="H081IE11.4"
Intron		/number=1 29436..29525
exon		/gene="H081IE11.4" /number=2 36680..36697 /gene="H081IE11.5"
exon		/number=1 join(36680..36697,37214..37384) /gene="H081IE11.5"
mRNA		36680..37384
gene		/gene="H081IE11.5" join(36680..36697,37214..37384) /gene="H081IE11.5"
CDS		/codon_start=1 /product="putative glucose-6-phosphate dehydrogenase" /protein_id="CAC09489.1" /db_xref="GI:10241651" /translation="MYMKLTFAAQIFPLTHDIDEGKKSIPYGSGRGPEADFSE RVGYMQTGHYIWIPPTLA" 36698..37213 /gene="H081IE11.5"
Intron		/number=1 37214..37384
exon		/gene="H081IE11.5" /number=2 complement(37968..38611) /gene="H081IE11.6"
exon		/number=3 complement(join(37968..38611,39696..39797,39902..39944)) /gene="H081IE11.6"
mRNA		complement(join(37968..38611,39696..39797,39902..39944)) /gene="H081IE11.6"
gene		complement(37968..39944) /gene="H081IE11.6"
CDS		complement(join(37968..38611,39696..39797,39902..39944))) /gene="H081IE11.6"
		/codon_start=1 /product="putative nucleotide-binding protein" /protein_id="CAC09490.1" /db_xref="GI:10241652"
		/translation="MENGCGDVPEANNDHCPTGSPAGAKADACTGCPNOICATAPK GPDVSIGSIFGLPDPDAIWMGRPKNGILKFELDVDWGIEDLYVDAPGTSDHI ISITQYLQINGIOGAIIVTTPPOVSLDIKKETINPKKYGVPIVLGVEMSLRQAFS DMKRYSEAGEEDATEMALNTIKRAPELLSVACSEVFDSKGEAEKMCQEMEVP LGKVPMPOLCKAEBERSCFYPDOKCSASAPALSKLIKLVTXK" complement(38612..39695) /gene="H081IE11.6"
Intron		/number=2 complement(39696..39797)
exon		/gene="H081IE11.6"
		/number=2 complement(39798..39901) /gene="H081IE11.6"
Intron		/number=1 complement(39902..39944) /gene="H081IE11.6"
exon		
Query Match	45.8%;	Score 106.8; DB 8; Length 59348;
Best Local Similarity	77.1%;	Pred. No. 8.4e-22;
Matches 145;	Conservative 0;	Mismatches 37; Indels 6; Gaps 1.
15	taggaatgttaaggcctcatagatatgggtggaagatttgagtatgatgaatgagcaa	74


```

Db 33703 TAAATATCTCAGTGGTTCTTCTATGACAGCGGAGAGATTGTAAGAAATTGAACAAGAA 33762
Oy 75 tctctgagatagagactcagaacaacatgactcggaaagctaccgcagatattatt 134
Db 33763 TTTCAGAGTATGAGAGTCA-----AACAAATCAGAAAGCCCTCGCAGGCTCTTATT 33816
Oy 135 tggcattgcctccatcagtcctaccatcagatgcgagatgataagatcattgcatga 194
Db 33817 TGGCATTGCTCCATCTGCTACTACCTTCAGTCTGCAAAATGATCAGAACATATTGTCATGA 33876
Oy 195 gtcacatc 202
Db 33877 ATCCATGT 33884

RESULT 5
OSJN00125
LOCUS OSJN00125 121690 bp DNA linear HTG 27-SEP-2001
DEFINITION Oryza sativa chromosome 4 clone OJ1672_A04, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
ACCESSION AL606998
VERSION AL606998.1 GI:15799239
KEYWORDS HTG; HTGS, PHASE2.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 121690)
Fu, G., Wang, S. Y., Ren, S. X., Lv, G., Lin, W., Gu, W. Q., Zhu, G. F.,
Tu, Y. F., Jia, J., Yin, H. F., Zhang, Y., Cai, Z., Chen, J., Kang, H.,
Chen, X. Y., Shao, C. Y., Sun, Y., Hu, Q. P., Zhang, X. L., Zhang, W.,
Wang, L. J., Ding, C. W., Sheng, H. H., Gu, J. L., Chen, S. T., Ni, L.,
Zhu, F. H., Han, B., Feng, Q., Huang, Y. C., Li, Y., Zhu, J. J., Zhao, Q.,
Hu, X., Liu, Y. L., Mu, J., Yu, Z., Chen, L., Fan, D. L., Meng, Q. J.,
Zhang, L., Lu, Y. Q., Yu, S. L., Liu, X. H., Lu, T. T., Zhang, Y. J., Lu, Y.,
Li, C., Li, T., Zhang, Y., Hu, H., Jia, P. X., Qian, Y. M., Ying, K.,
Zhou, B., Chen, Z. H., Hao, P., Zhang, L., Wu, M., Zhang, R. Q., Guan, J. P.
and Hong, G. F.
Direct Submission
Submitted (28-JUL-2000) Han Bin, National Center for Gene Research,
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
Oryza sativa japonica (niponbare) genomic DNA, chromosome 4, BAC
clone: OJ1672_A04.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Location/Qualifiers
1..121690
/organism="Oryza sativa"
/variety="Niponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="4"
/clone="OJ1672_A04"
/clone_11b="Monsanto-OJ"
BASE COUNT 34139 a 26001 c 26388 g 35144 t 18 others
ORIGIN
Query Match 45.8%; Score 106.8; DB 2; Length 121690;
Best Local Similarity 77.1%; Pred. No. 8,1e-22;
Matches 145; Conservative 0; Mismatches 37; Indels 6; Gaps 1;
Oy 15 tagaatattgtgctctatgataaggttggaagatttagtattgtaagaa 74
|||||

```

```

Db 62561 TAAATATCTCAGTGGTTCTTCTATGACAGCGGAGAGATTGTAAGAAATTGAACAAGAA 62620
Oy 75 tctctgagatagagactcagaacaacatgactcggaaagctaccgcagatattatt 134
Db 62621 TTTCAGAGTATGAGAGTCA-----AACAAATCAGAAAGCCCTCGCAGGCTCTTATT 62674
Oy 135 tggcattgcctccatcagtcctaccatcagatgcgagatgataagatcattgcatga 194
Db 62675 TGGCATTGCTCCATCTGCTACTACCTTCAGTCTGCAAAATGATCAGAACATATTGTCATGA 62734
Oy 195 gtcacatc 202
Db 62735 ATCCATGT 62742

RESULT 6
OSJN00053/c
LOCUS OSJN00053/c 151889 bp DNA linear HTG 11-SEP-2001
DEFINITION Oryza sativa chromosome 4 clone OSJNba0081L15, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
ACCESSION AL606623
VERSION AL606623.1 GI:15594082
KEYWORDS HTG; HTGS, PHASE2.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 151889)
Fu, G., Wang, S. Y., Ren, S. X., Lv, G., Lin, W., Gu, W. Q., Zhu, G. F.,
Tu, Y. F., Jia, J., Yin, H. F., Zhang, Y., Cai, Z., Chen, J., Kang, H.,
Chen, X. Y., Shao, C. Y., Sun, Y., Hu, Q. P., Zhang, X. L., Zhang, W.,
Wang, L. J., Ding, C. W., Sheng, H. H., Gu, J. L., Chen, S. T., Ni, L.,
Zhu, F. H., Han, B., Feng, Q., Huang, Y. C., Li, Y., Zhu, J. J., Zhao, Q.,
Hu, X., Liu, Y. L., Mu, J., Yu, Z., Chen, L., Fan, D. L., Meng, Q. J.,
Zhang, L., Lu, Y. Q., Yu, S. L., Liu, X. H., Lu, T. T., Zhang, Y. J., Lu, Y.,
Li, C., Li, T., Zhang, Y., Hu, H., Jia, P. X., Qian, Y. M., Ying, K.,
Zhou, B., Chen, Z. H., Hao, P., Zhang, L., Wu, M., Zhang, R. Q., Guan, J. P.
and Hong, G. F.
Direct Submission
Submitted (28-JUL-2000) Han Bin, National Center for Gene Research,
Chinese Academy of sciences, 500# Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: bhan@ncgr.ac.cn. Clone requests:
bhan@ncgr.ac.cn
Oryza sativa japonica (niponbare) genomic DNA, chromosome 4, BAC
clone: OSJNba0081L15.
Web site: http://www.ncgr.ac.cn
----- Summary Statistics
Assembly program: phrap
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Location/Qualifiers
1..151889
/organism="Oryza sativa"
/variety="Niponbare"
/sub_species="japonica"
/db_xref="taxon:4530"
/chromosome="4"
/clone="OSJNba0081L15"
/clone_11b="CUG1-OSJNba"
BASE COUNT 43345 a 32124 c 32302 g 44118 t
ORIGIN
Query Match 45.8%; Score 106.8; DB 2; Length 151889;
Best Local Similarity 77.1%; Pred. No. 8,1e-22;
Matches 145; Conservative 0; Mismatches 37; Indels 6; Gaps 1;
Oy 15 tagaatattgtgctctatgataaggttggaagatttagtattgtaagaa 74
|||||

```


Db 22636 TAAATATGTCAGTGGTCTCTATACAGCGGAGAGATTGTGAGAAATTGACAAAGGAAA 22577
QY 75 tctctgagatagagacttcagaaacaatgagactcggaagctaccagatattatc 134
Db 22576 TTTGAGATATGAGAGATCA-----AACAAATCAGAAAGAGCCCTCGAGGCTCTTCTATT 22523
QY 135 tggcatgtccctccatcagcttaccatcagatatacgagatgataagaatcattgtc 194
Db 22522 TGGCATTTGGCCCTCCATCTGTCTACCCCTTCAGTGTGCAAAATGATCAGAAATATTGATG 22463
QY 195 gtccatct 202
Db 22462 ATCCATGCT 22455

RESULT 7
MSU18238
LOCUS MSU18238 1785 bp mRNA linear PLN 30-JAN-1997
DEFINITION Medicago sativa glucose-6-phosphate dehydrogenase mRNA, complete
ACCESSION U18238
VERSION U18238.1 GI:603218
KEYWORDS alfalfa.
SOURCE Medicago sativa subsp. sativa
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago.

REFERENCE 1 (bases 1 to 1785)
AUTHORS Fahrendorf, T., Ni, W., Shorrosh, B.S. and Dixon, R.A.
TITLE Stresses responses in alfalfa (Medicago sativa L.) XX.
Transcriptional activation of oxidative pentose phosphate pathway genes at the onset of the isoflavonoid phytoalexin response
JOURNAL Plant Mol. Biol. 28 (5), 885-900 (1995)
MEDLINE 95367649
REFERENCE 2 (bases 1 to 1785)
AUTHORS Fahrendorf, T.
TITLE Direct Submission
JOURNAL Submitted (06-DEC-1994) Theo Fahrendorf, Samuel Roberts Noble Foundation, Plant Biology Division, 2510 Sam Noble Parkway, Ardmore, OK 73402, USA

FEATURES
source
Location/Qualifiers
1..1785
/organism="Medicago sativa subsp. sativa"
/cultivar="Apollo"
/sub_species="sativa"
/db_xref="taxon:56147"
/clone="G10"
/tissue_type="cell suspension culture"
38..1585
/codon_start=1
/product="glucose-6-phosphate dehydrogenase"
/protein_id="AA041552.1"
/db_xref="GI:603219"
/translation="MGTNEMHYERDSIGTESPVAREVLETGTLISIVLGASGLAKK
KTFPALFLYKOELPDPBEVHIFGARSKISDDELRLNLRSLVPEKASPKOLDLVK
KELQVYVSGPYSEDEGFLDKISFHEYLKKNKSSRLTFALAPSPVYSVCK
KITCCMKSKDGGWTRVYVEKPRERDLESABELSTOIGELPEEPQITRIDHYGKEL
VQNMVLRFANRFLPLMNNHIDVQVIFREDRETGDSGFQYGIIRDIIPNHL
OVCLIAEKPVSLKPEHIDREKVKVLESVLPIDREVDLQYEGYTDPPVPDSNT
PTFATTLIRINERWEGVPTVKAGKALNSKRAEIRVQFKDVGDIFFSKRGKNEFY
IRLOPSEAIYMKLTIVKQGLEMSAVQSELDLSYGRQYGTITPEAYERLLIDTIRGDO
OHFVRDELKASWQIFPLMLKIDRGELKPPYVNGSGRADELBELLEKAGYVOTPGI
WIIPYTL"

polyA_site
BASE COUNT 551 a 306 c 396 g 532 t
ORIGIN

* Query Match 45 8%; Score 106.6; DB 8; Length 1785;
Best Local Similarity 66.1%; Fred. No. 1.1e-21;

Matches 154; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 1 gttttgcagtgtagaataatgttggtgctccctatgataagggtggaagattgagtt 60
Db 346 GTTTTACCAATTTGTTAAATATGTAAGTGCCCTTATGTGATTTCTGAAGGATTTTCGCTT 405
QY 61 atgtaatgagcaatctctgatatgagacttcagaaacaatgagctcggaagctaccg 120
Db 406 GTTGGATTAAGAGATTTCAGAGCATGATATTGAAATAATAGTAAGAGGTTTCATCTCG 465
QY 121 cagattatttatttgcatctgctccatcagtcactacccatgatatgagatgataag 180
Db 466 GAGGCTTTCTATCTTGACATCTCTCCCTTCAGTGTATCCATCGTTTCAGAGATGATCA 525
QY 181 atcatattcagatgcatccttcacacacccggttgagaaaggttatgtgtg 233
Db 526 AACTTGTTCATGATTAATTAATCTGATCTTGTGTGATGAGACACCGCTTGTGTG 578

RESULT 8
STG6PDH
LOCUS STG6PDH 1689 bp mRNA linear PLN 21-SEP-1994
DEFINITION S. tuberosum mRNA for glucose-6-phosphate dehydrogenase.
ACCESSION X74421
VERSION X74421.1 GI:471344
KEYWORDS cytosolic enzyme; g6pdh gene; glucose-6-phosphate dehydrogenase; oxidative pentose phosphate pathway.
SOURCE potato.
ORGANISM Solanum tuberosum

REFERENCE 1 (bases 1 to 1689)
AUTHORS von Schaewen, A.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1993) A. von Schaewen, Universitaet Osnabrueck Pflanzenphys., Barbarastr 11, 49069 Osnabrueck, FRG
REFERENCE 2 (bases 1 to 1689)
AUTHORS Greave, K., von Schaewen, A. and Scheible, R.
TITLE Purification, characterization, and cDNA sequence of glucose-6-phosphate dehydrogenase from potato (Solanum tuberosum L.)
JOURNAL Plant J. 5 (3), 353-361 (1994)
MEDLINE 94236152

FEATURES
source
Location/Qualifiers
1..1689
/organism="Solanum tuberosum"
/cultivar="Desiree"
/db_xref="taxon:4113"
/clone="K4"
/haplotype="diploid"
/cell_line="S. tuberosum L. Desiree"
/tissue_type="mesophyll"
/tissue_type="green leaves"
/clone_lib="lambda ZAPit"
/dex_stage="vegetative"
46..1581
/gene="g6pdh"
46..1581
/gene="g6pdh"
/EC_number="1.1.1.49"
/note="homotetramer"
/codon_start=1
/product="glucose-6-phosphate 1-dehydrogenase"
/protein_id="CA52442.1"
/db_xref="GI:471345"
/db_xref="SWISS-PROT:P37830"
/translation="MAASKIEKRSIRINDSPFDNNINPETGCLSTIYVAGAGDLAKK
KTFPALFNLRGFLQSNVEYHIFGARTKISDDLSRIRGLYSQGENEGVSEFLQ
LIRYVSGSDSAAGFTSLDKAISHEHFKSKNSRRLFFALPSPVSPSCRMKLS
YCMKSKDLGGWTRVYVEKPRERDLESABELSTOIGELPEEPQITRIDHYGKELVNL
LVIRFANRFLPLMNNHIDVQVIFREDRETGDSGFQYGIIRDIIPNHLQVCLIAEKPV
SLKPEHIDREKVKVLESVLPIDREVDLQYEGYTDPPVPDSNTPTFATTLIRINERWEGV
PTVKAGKALNSKRAEIRVQFKDVGDIFFSKRGKNEFYIRLOPSEAIYMKLTIVKQGLEMSAVQSELDLSYGRQYGTITPEAYERLLIDTIRGDO
OHFVRDELKASWQIFPLMLKIDRGELKPPYVNGSGRADELBELLEKAGYVOTPGIWIIPYTL"

gene

CDS

`/db_xref="taxon:4043"`
`/cell_type="cultured cells"`
`1..1957`
`/gene="CG6PDH2"`
`74..1678`
`/gene="CG6PDH2"`
`/gene="CG6PDH2"`
`/codon_start=1`
`/product="cytosolic glucose-6-phosphate dehydrogenase 2"`
`/protein_id="AA66939.1"`
`/db_xref="GI:2352923"`
`/translation="MKSAVCDIVHKMSDKWCKLEEARPVEDSAGDPIVKTQPPYVS`
`EDCLSLIVLAGAGSDLARKKFPALFNLVYKQGFQSHSEVYIFGARIRISDDDLRDFI`
`RGLVLPKNTAEKTHEDVSKFIOLIKVCSYAAEGFCLDEISHEHSINIESS`
`SRRLFEYALPSPVYPVCWKIKTKYCNMKSQLGGWTRIVYKPRGRDLESAEHLNNG`
`ELPEEPOIVRIDHYKLEIYONLIVLIFEARFPLSNRNRTISVOYVPEEDCTEGR`
`GGYEDQGITRIDYIKNLLVYCLVAMERKISLKPENIRDEKKAIVDSVAPLNDEEV`
`LGQTDGLIDPITVPAGSNITPFAETWVIRINEREGVPEVLAKGALESRAELRIQF`

CDS

```

/cultivar="Samsun NN"
/db_xref="taxon:4097"
/cell_type="mesophyll"
/tissue_type="young sink leaves"
88..1623
/note="TCG9"
/codon_start=1
/product="glucose-6-phosphate dehydrogenase"
/protein_id="CA04993.1"
/db_xref="GI:3021510"
/db_xref="SPTREMBL:O65855"
/translation="MAASWCIEKRGSIKRLDFRNDNINIPETGSLITIVILGASDLAKK
KTPALNLYKROGFIOSNEVHIFGVARTKISDDILNGRIYLSOGKNEENEVESEFLQ
LTYVSGSYDSGEGEFSILDKAIEHATKSTGSSRRILFYFALPSVSPYRMKN
YCAKKSIDLGSMWTRIVYKPEFGKDLASEQSLQSIGELFPDRQIIRIDHLYGKRLVYL
LVAFKFAFRFLPLANNRIDNIDIVFREDFTGCGYFDEYGIKDIITONQLQVLC
LVMEKFRVQSPHEIRKREKVLQSMILPIKDEEIVLQEGYKRDPELVNNTSTPFA
```


BASE COUNT	515 a	340 c	434 g	532 t
ORIGIN	<p> TMTVLRINERNEGCVPTIMKACKALNSKAEIRVOFKVPGDIFCKCKÖGNEFVIRLQ PSEAMYNKLVYKRPGLDMSTVQSELDLSYRQYGVVIPLEVEYELLIDITIRGDOOHV RDLDAKAAEIEFTPLLRIDDEGKVPILPKNGSGRADEADELLÖNGVYÖTHGICIP PTL" </p>			

Query Match	41.6%	Score 97	DB 8	Length 1821
Best Local Similarity	63.5%	Pred No. 8.8e-19		
Matches 148	Conservative	0	Mismatches 85	Indels 0
			Gaps	0

OY	I	gtcttcgagtgatagaataatcgttagtgcgccatagatgaaggatcgaatt	60
Db	384	GTTCCTGCAGCGTGATTAAATACGTCAGGGCCCTTATATATCTGGGAGGCCCTTAGTTT	443
OY	61	atgaatgaagccaatccctctgatctlgagactcgaaacacaatgacctgggaagtaccg	120
Db	444	ACTGCACAAGCCTTAGCTGAGCACGAATAATGGCAAAAAATGACACAGAAGCATCATCCAG	503
OY	121	cagaattattttatcttgcaatgcctcccatcacgcttacccaatcagtatbcgaatgataag	180
Db	504	AAGACTCTTCTACTTTGGCTCTTCCTCCATCGTAATATCCCCTGGTTTTGCACAATGATPAA	563
OY	181	atcataatgcctgagtcacattctaacaacccgcttgacaagggtatctgtg	233
Db	564	AAATTACTGTTAGAACAAATCTGATCTGTGGTGGACCTCGCATCTTTCTGTGG	616

RESULT	11
LOCUS	AF012862
DEFINITION	1848 bp mRNA linear PLN 03-SEP-1997
ACCESSION	AF012862
VERSION	AF012862.1 GI:2352920
	Petrosselinum cytosolic glucose-6-phosphate dehydrogenase 1
	(cgo6pDHL mRNA, complete cds.

SOURCE	ORGANISM
parsley.	Petroselinum crispum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.	
1 (bases 1 to 1848)	
Batz,O., Logemann,E., Reinold,S. and Hahlbrock,K.	
Extensive reprogramming of cellular metabolism by fungal elicitor or infection in parsley suggests a new perception of 'defense-related' genes unpublished	

AUTHOR INDEX
 2 (pages 1 to 1848)
 AUTHORS
 Batz, O., Logemann, E. and Hallbrock, K.
 TITLE
 Direct Submission
 Submitted (08-JUL-1997) Blochemistry, MPI f. Zuechtungsforschung,
 Journal
 Carl-von-Linne-Weg 10, Cologne, NRW 50829, Germany
 FEATURES
 Location/Qualifiers

```

source
1..1848
/organism="Petroselinum crispum"
/db_xref="taxon:4043"
/cell_type="cultured cells"
1..1848
/gene="cg6PDH1"
41..1528
CDS

```

```

/gene="GG6PDH1"
/codon_start=1
/product="cytosolic glucose-6-phosphate dehydrogenase 1"
/protein_id="AA069318.1"
/db_xref="GI:2352921"
/translation="MTSDKWSIEKRAPLRNDSVLKENVPEIGCLSLIVLGSGLAK
KTFPALFNLYRQGLQSHSEVYIFGYATKLSIDEDLRIRICVLYSKNVEGHEDL
TNFIDLIKYCGSGYDEEGYOLDKREISEHEIYRNCQSGSRFLFLPLPSYAPVC
KMRKCSNKSNDLGGWTRIVAEKPKRGLFSAEQLSNOIGELFEPPQIYRIDTHYKLE
LYONLLVLFANRPMPMLNNDINIDIVFPREDGDTGRGQYFQOQYITINDITONL
LOVLYCAVMEKPYSLKLPKHIDEKVKVLAQSVPIKDESVYLGQYGLDEYLPDQSY
PTFTFTWLRILHNERWEGVPIKAKGKLNSKAKIRPQFQDVPDIDIKSKQGNF

```

BASE COUNT	590 a	318 c	409 g	531 t
ORIGIN	VIRLQPSSEAMYMKLTVKQPGKMKSTVOSEMDLISGGRYQDVPTTPEAYERLLIDTIRGD QQHVRBDEDLKAAWEITFTPLLRHTRDKDEKFSNLPYKPSGRPEESG"			

Query Match	41.2%	Score 96	DB 8	Length 1848
Best Local Similarity	63.4%	Pred. No.	1.8e-18	
Matches 147	Conservative	0	Mismatches 85	Indels 0
			Gaps	0

QY 2 tttttgagagtagtagaagatgtagtggcctccatagtagggtaggaagatttgagtta 61
 Db 353 TTTCTGCAGATTGATCAAAATATGTTGGCGATCTTATATGACCGGAGAGGGGTAAACAATT 412
 QY 62 ttgaatgaggaatcctcctcgagtatgagactcagaaacaatgactcggagagctaacgc 121
 Db 413 CTGGATTAAAGAAATATCTGAGCAGCAATATACAGAAATTTCACGGAAAGATCATCTCCGT 472
 QY 122 agatatactttatcttggcaatcgctccatccagttctaccatcagtatcgagatgataa 181
 Db 473 AGACATATTTTCCTGCTCTCTCCACATCAAGTATATATCTCTCTGTCAGATGATCAGG 532
 QY 182 tcatatgacagagtcacatctcaacaacacgggttgagcaagagtattctgtg 233
 Db 533 AAATGCTGACGATTAATATCTGATCTTGTTGGTGAAGAGCTCGGATGTTGTTGG 584

RESULT 12			
LOCUS	AY065054	1810 bp	mRNA linear
DEFINITION	Arabidopsis thaliana AT3g27300/K17E12.12 mRNA, complete cds.		
ACCESSION	AY065054		
VERSION	AY065054.1	GI:18086469	
KEYWORDS	FLI-CDNA.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		

REFERENCE
AUTHORS

Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots,
Rosidae: eurosids II: Brassicales; Brassicaceae; Arabidopsis
1 (passes 1 to 1810)

Cheuk, R., Chen, H., Kim, C. J., Koesema, E., Meyers, M. C., Banb, J.,
Bowers, L., Carninci, P., Chang, E., Dale, J. M., Goldsmith, A. D.,
Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Katlin-Nemmann, G.,
Kawai, J., Lam, B., Lee, J. M., Lin, J., Qi, M., Nakusaka, M.,
Nguyen, M., Onodera, C. S., Palm, C. J., March, H. L., Saurati, T.,
Satou, M., Seki, M., Southwick, A., Tang, C. C., Tortum, M., Wu, H. C.,
Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R. W.,
Theologis, A. and Ecker, J. R.

TITLE	Arabidopsis cDNA clones
JOURNAL	unpublished
REFERENCE	2 (bases 1 to 1810)
AUTHORS	Chen, R., Chen, H., Kim, C. J., Koesema, E., Meyers, M. C., Banh, J., Bowers, J., Carrinoci, P., Chang, E., Dale, J. M., Goldsmith, A. D., Hayshtlak, Y., Ishida, Y., Jones, T., Kamaya, A., Kallin-Neumann, G., Kawai, J., Lam, B., Lee, J. M., Lin, J., Miranda, M., Natusaka, M., Nguyen, M., Oonders, C. S., Palm, C. J., Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C. C., Tortoloni, M., Wu, H. C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.
TITLE	Direct Submission
JOURNAL	
REFERENCE	
AUTHORS	

COMMENT
The 2007 data include genomic analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salt Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA' : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamlya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shimozaki,K.

The Salk, Stanford, PCEC (SSP) Consortium members carried out the sequencing and annotation of the RAL1 cDNAs: Cheuk, R., Chen, H., Kim, C.-J., Koeseema, E., Meyers, M.C., Shinn, P., Banh, J., Bowser, L.,

Dale, J.M., Goldsmith, A.D., Jiang, P.X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Liu, S.X., Miranda, M., Nguyen, M., Onda, C.S., Palm, C.J., Pham, P.K., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

1. .1810

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/chromosome="3"

/clone="RAFL04-20-J16(R21172)"

/note="ecotype: Columbia"

1. .77

78. .1628

/note="putative glucose-6-phosphate 1-dehydrogenase"

/codon_start=1

/product="AT3g27300/K17E12.12"

/protein_id="AAL57688.1"

/db_xref="GI:18086470"

/translation="MSGQGMHEKRSILKNDSPYKENVPTETGSLIIVGASGDLAKKREPALFNFHOGFLNPDVHIFGVARSKITDDELNDKIRGIVLVDKNSKREALSKFLIKYVGPDPSEGRFRDLDAIIEHETSKTAGSSRLPYLPSVPPVSKMIRACWNKSDLGWTRIVKEPKGKDLSEASQIGALFEPPQYRIDHYLKELVQMMVLMEKRNFLPLMNRDNIANOIVREFPGTEGRGVEGYIIDIIONHLLOVCLVMEKRNFLPLMNRDNIANOIVREFPGTEGRGVEGYIIDIIONHLTPFATRTLRINRMWEGVPIFLKAGNSKADIRIOFGVPGDIFKCNOCNNEFVIRQPSBAMKLTIVKOPGLEMOTVSESLSKORODVSTPEATVIRLLDITRGDQOHVRDELKAAMEIFPLRLRIDGKSVKPSKSGRGAEPADQLKAGYMTGHIYIWPPL"

misc_difference

1204

/note="compared to genome"

/replace="a"

1628. .1810

3'UTR

BASE COUNT

538 a 322 c 416 g 534 t

ORIGIN

Query Match 40.9%; Score 95.4; DB 8; Length 1810;

Best Local Similarity 63.1%; Pred. No. 2.7e-18;

Matches 147; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1 gtttttcagttagtagaataatgtagtgcctcctatgataagtagtggaagatttgagt 60

DB 389 GTTCTTAAGCGTAAAGTAAATGATAGTACGACCTTATGATTTCTGAGAGCGTTTAAAG 448

QY 61 attgaatgagcaatcctcgaatgtagatgagcagaacaaatgactcggaagctaccg 120

DB 449 ATTGACAAAGCGGATTTGGACGACAGATATCTAAAGAAAGACTGCGAGGATCTTCGAG 508

QY 121 cagatattatt 180

DB 509 GAGATGTTTATCTTGACCTTCCTCGCTGTATACCTCTCTGTAAAGCAAAATGATCA 568

QY 181 atcatattgacagatgcatccttcacacccggttgacgaaggttatgtgtc 233

DB 569 GGCAATGGTCACCTAACAAATCTGATCTTGATGATGAGATAGGATTTGTTG 621

RESULT 13

NTTCG6

LOCUS 1852 bp mRNA linear PLN 17-AUG-1999

DEFINITION Nicotiana tabacum mRNA cytosolic glucose-6-phosphate dehydrogenase

ACCESSION AJ001769

VERSION AJ001769.1 GI:3021507

KEYWORDS common tobacco.

SOURCE Nicotiana tabacum

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots: Asteridae: euasterids I; Solanales; Solanaceae; Nicotiana. (bases 1 to 1852)

Wendt, U.K., Hauschild, R., Lange, C., Pietersma, M., Wenderoth, I. and von Schaewen, A.

Evidence for functional convergence of redox regulation in G6PDH

isoforms of cyanobacteria and higher plants

Plant Mol. Biol. 40 (3), 487-494 (1999)

2 (bases 1 to 1852)

von Schaewen, A.

Direct Submission

Submitted (18-SEP-1997) von Schaewen A., Plant Physiology,

University Osnabrueck, Barbarastr. 11, D-49069 Osnabrueck, GERMANY

Location/Qualifiers

1. .1852

/organism="Nicotiana tabacum"

/cultivar="Samsun NN"

/db_xref="taxon:4097"

/cell_type="mesophyll"

/tissue_type="young sink leaves"

102. .1634

/note="TCG6"

/codon_start=1

/product="glucose-6-phosphate dehydrogenase"

/protein_id="CAA04992.1"

/db_xref="GI:3021508"

/translation="MAASMCIEKRSILKNDSPRENDNIPEITGSLIIVGASGDLAKKREPALFNFHOGFLNPDVHIFGVARSKITDDELNDKIRGIVLVDKNSKREALSKFLIKYVGPDPSEGRFRDLDAIIEHETSKTAGSSRLPYLPSVPPVSKMIRACWNKSDLGWTRIVKEPKGKDLSEASQIGALFEPPQYRIDHYLKELVQMMVLMEKRNFLPLMNRDNIANOIVREFPGTEGRGVEGYIIDIIONHLLOVCLVMEKRNFLPLMNRDNIANOIVREFPGTEGRGVEGYIIDIIONHLTPFATRTLRINRMWEGVPIFLKAGNSKADIRIOFGVPGDIFKCNOCNNEFVIRQPSBAMKLTIVKOPGLEMOTVSESLSKORODVSTPEATVIRLLDITRGDQOHVRDELKAAMEIFPLRLRIDGKSVKPSKSGRGAEPADQLKAGYMTGHIYIWPPL"

BASE COUNT 525 a 353 c 432 g 542 t

ORIGIN

Query Match 40.9%; Score 95.4; DB 8; Length 1852;

Best Local Similarity 63.1%; Pred. No. 2.7e-18;

Matches 147; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 1 gtttttcagttagtagaataatgtagtgcctcctatgataagtagtggaagatttgagt 60

DB 395 GTTCTGCAACTGATTAATATACGTCAGTGGCTTATGATTTCTGAGAGCGCTTTAGTTT 454

QY 61 attgaatgagcaatcctcgaatgtagatgagcagaacaaatgactcggaagctaccg 120

DB 455 ACTGACAAAGCGGATATGTCGACGACGAATTCGCAAAAATATAGACGAGGATCATCCAG 514

QY 121 cagatattatt 180

DB 515 AAGACTCTTCTACTTTCCTTCCTCCATCAGATATTCCTCTGTTTGCAGAGATGATAA 574

QY 181 atcatattgacagatgcatccttcacacccggttgacgaaggttatgtgtc 233

DB 575 AAACATATGTATGACAAATCTGATCTTGATGATGAGATAGGATTTGTTG 627

RESULT 14

ATH010970

LOCUS 1862 bp mRNA linear PLN 17-AUG-1999

DEFINITION Arabidopsis thaliana mRNA for cytosolic glucose-6-phosphate

1-dehydrogenase AC99.

ACCESSION AJ010970

VERSION AJ010970.1 GI:5732194

KEYWORDS acg9 gene; glucose-6-phosphate 1-dehydrogenase.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 01:10:45 ; Search time 796.36 Seconds

(without alignments)
502.337 Million cell updates/sec

Title: US-09-300-482-1

Perfect score: 233

Sequence: 1 gtttcgactgtagtaat.....ttgacgaagggtattgtt 233

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_032802:*

1: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	97	41.6	1915	21	AAC47923	Arabidopsis thaliana
2	34.6	14.8	597	22	AAH52396	S. epidermidis open
3	34.6	14.8	2945	22	AAH54407	S. epidermidis genome
4	34.6	14.8	3825	22	AAH54430	S. epidermidis genome
5	34.4	14.8	10286	22	AAH545308	Chemically pretreated
6	31.8	13.6	621	22	AAH31486	Human olfactory receptor
7	31.8	13.6	621	22	AAH32502	Human olfactory receptor
8	31.6	13.6	278	21	AAA01024	Human colon cancer
9	31.4	13.5	13376	24	ABJ32583	Human immune system

10	31.2	13.4	4418	18	AAV74765	Staphylococcus aur
11	31	13.3	54863	22	AAK86025	Human immune/haema
12	31	13.3	54877	22	AAK86026	Human immune/haema
13	30.6	13.1	661	23	AAH71803	DNA encoding novel
14	30.6	13.1	925	22	AAI97683	Human neuroblastom
15	30.6	13.1	2322	24	ABJ99851	Mouse ischaemic co
16	30.6	13.1	3046	21	AAH78092	Human cancer assoc
17	30.6	13.1	4237	21	AAH78092	Human cancer assoc
18	30.6	13.1	12468	22	AAK95240	Human immune/haema
19	30.6	13.1	1503900	22	AAK95240	Human immune/haema
20	30.6	13.1	1503900	22	AAK95240	Human immune/haema
21	30.4	13.0	60747	23	ABJ16128	Drosophila melanog
22	30	12.9	11812	22	AAH45502	Chemically pretrea
23	30	12.9	11812	22	AAH45502	Chemically pretrea
24	30	12.9	11812	22	AAH45502	Chemically pretrea
25	30	12.9	12054	24	ABJ33179	Human immune syste
26	29.8	12.8	5280	18	AAV74855	Staphylococcus aur
27	29.8	12.8	5631	20	AAH85506	DNA encoding. Sac
28	29.8	12.8	6920	22	AAH24300	Legionella pneumop
29	29.8	12.8	7655	16	AAH86896	S. cerevisiae fks1
30	29.8	12.8	26997	22	AAH46747	Tumour suppressor
31	29.8	12.8	513445	22	AAH161373	Soybean 318013 reg
32	29.6	12.7	13788	22	ABJ19999	Human nervous syst
33	29.4	12.6	786	22	AAH66049	C glutamicum codin
34	29.4	12.6	849	22	AAH71610	Corynebacterium g1
35	29.4	12.6	3271	23	ABJ03784	Drosophila melanog
36	29.4	12.6	3861	23	ABJ03784	Drosophila melanog
37	29.4	12.6	7810	22	AAH45434	Chemically pretrea
38	29.4	12.6	14335	22	AAH28524	Genomic fragment #
39	29.4	12.6	44014	23	ABJ02654	Drosophila melanog
40	29.4	12.6	349980	22	AAH68557	C glutamicum codin
41	29.2	12.5	687	22	AAH32558	Human genomic DNA
42	29.2	12.5	687	22	AAH32558	Human genomic DNA
43	29.2	12.5	687	22	AAH32558	Human genomic DNA
44	29.2	12.5	687	22	AAH32558	Human genomic DNA
45	29.2	12.5	687	22	AAH32558	Human genomic DNA

ALIGNMENTS

RESULT 1

AAC47923 standard; DNA: 1915 BP.

AC AAC47923:

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 55610.

XX

XX

XX Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

XX

OS Arabidopsis thaliana.

OS

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137232.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140961.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.


```
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159363.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 41.6%; Score 97; DB 21; Length 1915;
Best Local Similarity 63.5%; Pred. No. 2.3e-22;
Matches 148; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```
OY 1 gttttgcagtgatgaatagttagtgcttcctatgataggtggaaggtattgatt 60
    ||||| || || || || || || || || || || || || || || || || ||
DB 412 gtctctaaagcgcgattaaatattgttgaccttattgtcgaaggaaggttaag 471
    || || || || || || || || || || || || || || || || || || ||
OY 61 attgaatggagcaatctcgaatgatgacctcagaagaacaatgacctcggaagctaccg 120
    || || || || || || || || || || || || || || || || || || ||
DB 472 attagaacaaggagatttcggacacgagataatcctaaagaagctgcgaaggtactcttag 531
    || || || || || || || || || || || || || || || || || || ||
OY 121 cagatattatttttgcatcttgccatcgcacatcagtcacatcagatagcgagatgaag 180
    || || || || || || || || || || || || || || || || || || ||
DB 532 gagatgttttattcttcgacttcctcgcgtctgtatataccctcctgaagaagatgatcaa 591
    || || || || || || || || || || || || || || || || || || ||
OY 181 atccatattcagatgcattcattcaacacccggttgacaaggttatgttg 233
    || || || || || || || || || || || || || || || || || || ||
DB 592 ggcatgtgtgcactacaacatctgattctgtgtgagtactaagattgttctgtg 644
    || || || || || || || || || || || || || || || || || || ||
```

RESULT 2

AAH52396
ID AAH52396 standard; DNA; 597 BP.

XX AAH52396;

DT 03-SEP-2001 (first entry)

XX S. epidermidis open reading frame nucleotide sequence SEQ ID NO:185.

XX Staphylococcus epidermidis SRI strain; infection; diagnosis;

KW vaccination; endocarditis; ds.

XX Staphylococcus epidermidis.

OS WO200134809-A2.

XX WO200134809-A2.

XX 17-MAY-2001.

PD 09-NOV-2000; 2000WO-US30782.

XX 09-NOV-1999; 99US-0164258.

XX (GLAX) GLAXO GROUP LTD.

XX PA

XX PI

XX Kimmerly WJ;

XX DR

DR WPI; 2001-316495/33.
DR P-PSDB; AAG81546.
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT useful for vaccinating against infections, e.g. endocarditis -
PS Claim 8; Page 93-94; 2188pp; English.

XX AAH53904 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81546 to AAG83120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH5091 to
CC AAH5098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.

XX Sequence 597 BP; 230 A; 101 C; 87 G; 179 T; 0 other;

XX Query Match 14.8%; Score 34.6; DB 22; Length 597;

XX Best Local Similarity 61.8%; Pred. No. 0.13;
Matches 55; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

OY 11 ttatagaatatgttgcttcctatgataggttggaaggtattgattatgattgag 70

DB 436 ttagaagaatttaattatgaccacgaagatgactggaagatttaattatgaaatcaa 495

OY 71 gcaatcttgatgattcagaaga 99

DB 496 cttaacctgattatcaaacaccagaaga 524

RESULT 3

AAH54007/C
ID AAH54007 standard; DNA; 2945 BP.

XX AAH54007;

DT 03-SEP-2001 (first entry)

XX S. epidermidis genomic polynucleotide sequence SEQ ID NO:3371.

XX Staphylococcus epidermidis SRI strain; infection; diagnosis;

KW vaccination; endocarditis; ds.

XX Staphylococcus epidermidis.

OS WO200134809-A2.

XX WO200134809-A2.

XX 17-MAY-2001.

PD 09-NOV-2000; 2000WO-US30782.

XX 09-NOV-1999; 99US-0164258.

XX (GLAX) GLAXO GROUP LTD.

XX PA

XX PI

XX Kimmerly WJ;

XX WPI; 2001-316495/33.

XX DR

XX	Nucleic acids encoding polypeptides from <i>Staphylococcus epidermidis</i> , useful for vaccinating against infections, e.g. endocarditis - Claim 8; Page 921-922; 2188pp; English.
PT	
XX	
XX	AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides (II), given in AAG61454 to AAG83120, from <i>Staphylococcus epidermidis</i> . CC (I) and (II) can have antibacterial activity and therefore can be used CC in vaccination. The nucleic acids (I) may be used to produce the CC S. epidermidis polypeptides (II) via the production of vectors CC containing them which are used to produce host cells which express the CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be CC used to vaccinate subjects and to raise antibodies against the bacteria. CC The polypeptides may also be used to assay for other inhibitors of their CC activity and therefore identify compounds that may be used for the CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to CC AAH55090 represent specifically claimed S. epidermidis genomic DNA CC polynucleotide sequences from the present invention. AAH5091 to CC in the exemplification of the present invention. CC N.B. The present invention specifically claims all the polynucleotide CC sequences given in the sequence listing of the present specification, CC however the sequence listing only goes up to SEQ ID NO:4454 so even CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472, CC no sequences are present for SEQ ID NO:4455 to 4464.
XX	
XX	Sequence 2945 BP; 1064 A; 436 C; 484 G; 961 T; 0 other;
XX	
XX	Query Match 14.8%; Score 34.6; DB 22; Length 2945; Best Local Similarity 61.8%; Pred. No. 0.25; Matches 55; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY	11 ttatgaataatgttgtagctcctatgtagtgaggtgaagattgagttatgaatgag 70 DB 162 TTGAGAGATTATTATATGACCGACGAGATCTGGGCAATTATTAATGTATGAATCA 103 QY 71 gcaatcttgagtatgagacttcaagaana 99 DB 102 CTAACTCTGATTTATCAACCCGAGAGA 74
DB	
XX	
XX	RESULT 4 AAH54430/c ID AAH54430 standard; DNA; 3825 BP.
XX	
XX	AAH54430;
XX	
XX	03-SEP-2001 (first entry)
XX	
XX	S. epidermidis genomic polynucleotide sequence SEQ ID NO:3794.
DE	
XX	Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination; endocarditis; ds.
KW	
XX	Staphylococcus epidermidis.
OS	
XX	MO200134809-A2.
PN	
PD	17-MAY-2001.
XX	
XX	09-NOV-2000; 2000WO-US30782.
PF	
XX	09-NOV-1999; 99US-0164258.
PR	
XX	(GLAX) GLAXO GROUP LTD.
PA	
XX	Kimmerly WJ;
PI	
XX	WPI; 2001-316495/33.
DR	
XX	
XX	Nucleic acids encoding polypeptides from <i>Staphylococcus epidermidis</i> ,

useful for vaccinating against infections, e.g. endocarditis -

Claim 8: Page 1427-1428: 2188bp; English.

AAH53970 represent nucleic acids (I) encoding polypeptides (II), given in AMG81454 to AMG83120, from *Staphylococcus epidermidis*. CC (I) and (II) can have antibacterial activity and therefore can be used in vaccination. The nucleic acids (I) may be used to produce the S. epidermidis polypeptides (II) via the production of vectors containing them which are used to produce hosts cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of their activity and therefore identify compounds that may be used for the treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to AAH55090 represent specifically claimed S. epidermidis genomic DNA polynucleotide sequences from the present invention. AAH55091 to CC in the exemplification of the present invention.

N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequence listing only goes up to SEQ ID NO:4454 so even though sequences are given in the disclosure for SEQ ID NO:4455 to 4472, no sequences are present for SEQ ID NO:4455 to 4464.

Sequence 3825 BP; 1407 A; 526 C; 643 G; 1249 T; 0 other;

Query Match 14.8%; Score 34.6; DB 22; Length 3825; Best Local Similarity 61.8%; Pred. No. 0.28; Matches 55; Conservative 0; Mismatches 34; Indels 0; Gaps 0

11 ttatagaataatgttagtgacctctatgaagggtggaagatttgatttcaatgag 70
||||| ||| | ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
780 TTAGAAGATTATTTATATGACCAAGAGTACTGGCAATTATTAAGTTATGATCAA 721
|| ||||| ||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
71 gaactcttgattgtagacttcagaaa 99
|| ||||| ||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
720 CTAACCTGTGATTATCAACCCAGAGA 692

RESULT 5
AA545308 standard; DNA; 10286 BP.

AA545308: 18-DEC-2001 (first entry)

Chemically pretreated genomic DNA associated with cell cycle #7.

Cell cycle; human; CPG dinucleotide; cytosine methylation; HIV; aging; human immunodeficiency virus; neurodegenerative disorder; solid tumour; graft-versus-host disease; glomerular disease; Lewy body disease; cancer; arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic; immunosuppressive; antitumour; cyostatic; antiarteriosclerotic; ds; PCR primer.

Homo sapiens.

WO200168911-A2.

20-SEP-2001.

15-MAR-2001; 2001WO-EP02945.

15-MAR-2000; 2000DE-1013847.
06-APR-2000; 2000DE-1018058.
07-APR-2000; 2000DE-1019173.
30-JUN-2000; 2000DE-1032529.
01-SEP-2000; 2000DE-1043826.

(EPig-) EPIGENOMICS AG.


```

XX XX New polynucleotides which encode polypeptides involved in olfactory
PT PT sensation for identifying olfactory agonists and antagonists -
XX XX
XX PS Claim 8; Page 611; 1857pp; English.
XX XX
CC CC The present sequence is one of a number of isolated polynucleotides
CC CC which encode polypeptides involved in olfactory sensation. The
CC CC polynucleotides can be used in screening for olfactory agonists and
CC CC antagonists. The methods allow for the determination of primary
CC CC scents and the identification of the odour receptors used to detect
CC CC secondary scents. The methods also enable determination of
CC CC receptors that are involved in detecting such secondary scents.
CC CC This enables the construction of a scent representation (also called
CC CC a scent fingerprint or scent profile), which may be used to re-create
CC CC and edit scents. Libraries of olfactory receptors are useful for
CC CC determining the interaction pattern of a composition with the receptors,
CC CC and can be used for determining differences in the olfactory faculties
CC CC of different individuals.
XX XX
SQ Sequence 621 BP; 212 A; 85 C; 180 G; 136 T; 8 other:

Query Match      13.6%; Score 31.8; DB 22; Length 621;
Best Local Similarity 55.6%; Pred. No. 1.2;
Matches 60; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 53 tttagatattgaatgagcaatctcagatgacttcagaagaacatgactggga 112
   || || || || || || || || || || || || || || || || || || ||
Db 511 TGTGATATCTCTCTTGGCGTCTCATACACTCTGACTTCAGTACACATGACATGCCA 452
   || || || || || || || || || || || || || || || || || || ||

QY 113 agctacgcgcagatattatttattggcattgcctccatcagctaccca 160
   || || || || || || || || || || || || || || || || || || ||
Db 451 GGGATCCCACTGTATGTCTCTACCCATGCGCTCTCCCTTGCCACCCA 404
   || || || || || || || || || || || || || || || || || || ||

RESULT 8
AAA01024
ID AAA01024 standard; cDNA; 278 BP.
XX XX
AC AAA01024;
XX XX
DT 19-MAY-2000 (first entry)
XX XX
DE Human colon cancer cell line polynucleotide sequence SEQ ID NO:1015.
XX XX
KW Human; colon cancer; tumour; diagnosis; gene expression product;
KW Probe; detection; cancerous state; metastasis; identification;
KW breast cancer; oestrogen receptor-positive breast cancer; therapy;
KW oestrogen receptor-negative breast cancer; lung cancer; ss.
XX XX
OS Homo sapiens.
XX XX
PN WO958675-A2.
XX XX
PD 18-NOV-1999.
XX XX
PF 13-MAY-1999; 99WO-US10602.
XX XX
PR 14-MAY-1998; 98US-0085426.
PR 15-MAY-1998; 98US-0085537.
PR 15-MAY-1998; 98US-0085696.
PR 21-OCT-1998; 98US-0105234.
PR 27-OCT-1998; 98US-0105877.
XX XX
PA (CHIR ) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J,
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A,
PI Lamsan G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I,
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B,

```

```

XX XX WPI; 2000-126369/11.
DR DR Polynucleotide library used to determine cancerous states of mammalian
XX XX cells -
PT PT
XX XX
XX PS Claim 1; Page 426; 1097pp; English.
XX XX
CC AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
CC libraries constructed from human colon cancer cell lines. The present
CC invention also describes a method of detecting differentially expressed
CC genes correlated with a cancerous state of a mammalian cell comprising
CC detecting at least one differentially expressed gene product in a test
CC sample derived from a cell suspected of being cancerous, where detection
CC of the differentially expressed gene product is correlated with a
CC cancerous state of the cell from which the test sample was derived.
CC The polynucleotides sequences can be used in a method for detecting
CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of
CC pre-metastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer.
XX XX
SQ Sequence 278 BP; 73 A; 49 C; 49 G; 78 T; 29 other:

Query Match      13.6%; Score 31.6; DB 21; Length 278;
Best Local Similarity 56.9%; Pred. No. 0.97;
Matches 58; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 34 ctatgtagggtggaaggattgagttatgaaatgagcaactctgagtagaacttc 93
   || || || || || || || || || || || || || || || || || || ||
Db 48 ctacagtaagaatgctatttatttatttatttatttatttatttatttattt 107
   || || || || || || || || || || || || || || || || || || ||

QY 94 agaaacaatgactcggaagctaccgacagattatttattt 135
   || || || || || || || || || || || || || || || || || || ||
Db 108 accagatgatgacaagggttaacaggtacttatttattt 149
   || || || || || || || || || || || || || || || || || || ||

RESULT 9
ABL32583
ID ABL32583 standard; DNA; 13376 BP.
XX XX
AC ABL32583;
XX XX
DT 26-MAR-2002 (first entry)
XX XX
DE Human immune system associated gene SEQ ID NO: 556.
XX XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antiparasitic;
KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX XX
OS Homo sapiens.
XX XX
PN WO200200928-A2.
XX XX
PD 03-JAN-2002.
XX XX
PF 02-JUL-2001; 2001WO-EP07537.
XX XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX XX
PA (EPIC-) EPIDENOMICS AG.

```


XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation
XX
XX Claim 1; SEQ ID NO 556; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
XX Sequence 13376 BP; 4172 A; 190 C; 2688 G; 6324 T; 2 other;

Query Match 13.5%; Score 31.4; DB 24; Length 13376;
Best Local Similarity 49.7%; Pred. No. 5.4;

Matches 80; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 54 ttgagttatgaaatgagcaatctctgagatgagactcagaacaaacatgctgggaa 113
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 5427 tttaaatattgattgagattatttaagaagataagattgttattagatttggggag 5486

QY 114 gctacgcgagatattattattgcatgctccatcagctcaccatcagatgagaga 173
|| || || || || || || || || || || || || || || || || || || || ||

DB 5487 tgtaaagacgagtgatattgtgtgtttgaaatgattgttggtaggagagt 5546

QY 174 tgaatgacatattgcatgagcattcaccacaccggt 214
|| || || || || || || || || || || || || || || || || || || || ||

DB 5547 tttttaaaaaagtaagaagttaattatcgtagatagtt 5587

RESULT 10
AAV74765
AAV74765 standard; DNA; 4418 BP.

XX AAV74765;

XX 16-MAR-1999 (first entry)

XX Staphylococcus aureus contig SEQ ID #454.

XX Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KM skin infection; surgical wound infection; scalded skin syndrome;
KM toxic shock syndrome; ds.

XX Staphylococcus aureus.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

the sequence listing in the specification. They
are included to maintain the nucleotide numbering
given in the specification for this DNA sequence

FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence
XX
XX EP786519-A2.
XX
XX 30-JUL-1997.
XX
XX 07-JAN-1997; 97EP-0100117.
XX
XX 05-JAN-1996; 96US-0009861.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
XX Rosen CA;
XX WPI; 1997-374922/35.
XX
XX Polynucleotide(s) and proteins derived from Staphylococcus aureus
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
XX
XX Claim 1; Page 1368-1371; 3271pp; English.

CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC S.aureus DNA sequences can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the
CC computer readable medium.
XX
XX

SEQ Sequence 4418 BP; 1109 A; 736 C; 658 G; 1732 T; 183 other;

Query Match 13.4%; Score 31.2; DB 18; Length 4418;
Best Local Similarity 53.2%; Pred. No. 4;

Matches 66; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 48 aagatttgattatgaatgagcaatctctgagatgagactcagaacaaacatgact 107
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 1312 aaagattttgtttatattatcttggaattcttaacagaaacacgcgacattgacc 1371

QY 108 cgggaagctaccgcagatattattattggaatgctccatcagatcagatcagat 167
|| || || || || || || || || || || || || || || || || || || || ||

DB 1372 tggtaagtaatacatattattattgtttttacgacgacgaaataacaataatgat 1431

QY 168 gcga 171

DB 1432 ccta 1435

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX Human: immune: haematopoietic: immune/haematopoietic antigen: cancer:
KW Cytostatic: gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225265.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227809.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229519.
PR 06-SEP-2000; 2000US-0229517.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0235837.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249269.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251858.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.

PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0251989.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251859.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-0255678.
 XX
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-483426/52.
 XX
 PT

Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 useful for preventing, diagnosing and/or treating cancers and

PT metastasis -
 XX
 PS Disclosure: SEQ ID NO 40838; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 54877 BP; 15008 A; 10329 C; 11472 G; 18068 T; 0 other;

Query Match 13.3%; Score 31; DB 22; Length 54877;
 Best Local Similarity 52.8%; Pred. No. 13;
 Matches 67; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 16 agaataatgttagtggtcctcctatgatagggtggaagattggtatgaagagcaat 75
 Db 12463 aaattacacttggtgcctacatgatattccttggatagaagcattagaccacatt 12522
 QY 76 ctctgagatgagcttcgaaacacatgactcgcggaagctcgcagattatttatt 135
 Db 12523 ctgagagcttaagacagatgatttgaagacactgagagacttatttatttatt 12582
 QY 136 ggcattg 142
 Db 12583 gtttttg 12589

RESULT 13
 AAS71803
 ID AAS71803 standard; cDNA; 661 BP.
 XX
 AC AAS71803;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #7607.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PE 30-MAR-2001; 2001WO-US08631.
 XX
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG07616.
 XX

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

PS Claim 1: SEQ ID No 7607; 103pp; English.

XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AA564197-AA594564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 661 BP; 175 A; 188 C; 97 G; 201 T; 0 other;

Query Match 13.1%; Score 30.6; DB 23; Length 661;

Best Local Similarity 56.4%; Pred. No. 3;

Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 112 aagctacgcgaattatttattgcatgctccatcagtcctaccatcagtagcga 171

DB 330 aggcctacgcgaattatttattgcatgctccatcagtcctaccatcagtagcga 389

QY 172 gatgataagatcatattgcatgagtcacatcttcacacaccg 212

DB 390 ccttaccaccaaatcttcttcagtgatgacatctcctccactg 430

RESULT 14

AA197683

ID AA197683 standard; cDNA; 925 BP.

XX AA197683;

XX 13-NOV-2001 (first entry)

DE Human neuroblastoma expressed polynucleotide SEQ ID NO 3758.

KM Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.

XX Homo sapiens.

XX WO200166719-A1.

XX 13-SEP-2001.

XX 02-MAR-2001; 2001WO-JP01629.

XX 07-MAR-2000; 2000JP-0159195.

PA (CHIB-) CHIBA PREPARECTURE.

PA (HISM) HISAMITSU PHARM CO LTD.

XX Nakagawara A;

XX WPI; 2001-565584/63.

XX Nucleic acids originating in gene expressed in human neuroblastoma,
 PT useful as probe or primer in diagnosing prognosis of human
 PT neuroblastoma, malignancy and susceptibility indicator or tumour marker
 PT for anti-cancer agents

PS Claim 1: Page 2740; 2979pp; Japanese.

XX
 CC The invention relates to novel genes (AA193926-AA197963) expressed in
 CC human neuroblastoma. The nucleic acids are applicable as a probe or
 CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and
 CC susceptibility indicators or tumour markers for anti-cancer agents. The
 CC gene information for diagnosing prognosis is related to factors similar
 CC to that for N-myc and TrkA genes.

XX Sequence 925 BP; 273 A; 149 C; 189 G; 287 T; 27 other;

Query Match 13.1%; Score 30.6; DB 22; Length 925;

Best Local Similarity 56.4%; Pred. No. 3.4;

Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 111 gaagctacgcgaattatttattgcatgctccatcagtcctaccatcagtagc 170

DB 392 gaagctgaagaagtgatcgaatttgaactaccaccctcagaagaagcagaggtg 451

QY 171 agatgataagatcatattgcatgagtcacatcttcacacacc 211

DB 452 tagataaatgcatgctcctcagtgctcctggtctcttaacc 492

RESULT 15

AB199851

ID AB199851 standard; cDNA; 2322 BP.

XX AB199851;

XX 07-MAR-2002 (first entry)

DE Mouse ischemic condition related cDNA sequence SEQ ID NO:979.

KM Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;

KM vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.

XX Mus musculus.

XX WO200188188-A2.

XX 22-NOV-2001.

XX 18-MAY-2001; 2001WO-JP04192.

XX 18-MAY-2000; 2000JP-0145977.

PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

XX WPI; 2002-034733/04.

XX P-PSDB; ABB57349.

XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or
 PT by determining the expression profile of a gene group comprising these
 PT genes

PS Claim 2: Page 2467-2471; 2690pp; English.

XX The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischaemic condition (e.g. compressive

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 01:00:35 ; Search time 185.98 Seconds
(without alignments)
307.735 Million cell updates/sec

Title: US-09-300-482-1

Perfect score: 233
Sequence: 1 gttttcagttagtagaat.....ttggacaaggttatgtttg 233

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.NA.*

1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/PCRTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/Backfilest.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	29.8	12.8	7655	1	US-08-619-554-1 Sequence 1, Appli
C 2	28.4	12.2	2596	4	US-09-289-254-1 Sequence 1, Appli
C 3	28.4	12.2	17710	4	US-08-976-259-70 Sequence 70, Appli
4	28.2	12.1	12565	4	US-09-345-217-3 Sequence 3, Appli
5	28	12.0	3417	2	US-08-464-402-1 Sequence 1, Appli
6	28	12.0	3417	4	US-09-054-775C-1 Sequence 1, Appli
7	27.4	11.8	527	4	US-08-976-259-137 Sequence 137, App
8	27.4	11.8	1172	4	US-08-936-165A-8 Sequence 8, Appli
9	27.4	11.8	2442	3	US-09-040-485-1 Sequence 1, Appli
10	27.4	11.8	5892	3	US-08-755-587-27 Sequence 27, Appli
11	27.4	11.8	7240	3	US-08-755-587-15 Sequence 15, Appli
12	27.4	11.8	11283	2	US-08-603-753D-3 Sequence 3, Appli
13	27.4	11.8	11283	3	US-09-099-753-3 Sequence 3, Appli
14	27.4	11.8	11283	4	US-08-986-106-3 Sequence 3, Appli
15	27.4	11.8	11385	2	US-08-639-501-1 Sequence 1, Appli
16	27.4	11.8	11385	3	US-09-044-946-1 Sequence 1, Appli
17	27.4	11.8	11385	3	US-09-044-908-1 Sequence 1, Appli
C 18	27.2	11.7	1153	4	US-08-858-207A-126 Sequence 126, App
C 19	26.6	11.4	3411	4	US-08-146-969-4 Sequence 4, Appli
20	26.6	11.4	4656	4	US-09-425-665-1 Sequence 1, Appli
21	26.6	11.4	4656	4	US-09-685-668-1 Sequence 1, Appli
C 22	26.4	11.3	1791	1	US-08-245-294-7 Sequence 7, Appli
C 23	26.4	11.3	1791	1	US-08-474-499-7 Sequence 7, Appli
C 24	26.4	11.3	1791	1	US-08-307-279A-7 Sequence 7, Appli
C 25	26.4	11.3	1791	5	PCT-US95-06211-7 Sequence 7, Appli
26	26.4	11.3	1917	3	US-08-755-587-1 Sequence 1, Appli
C 27	26.2	11.2	3159	3	US-08-986-485-3 Sequence 3, Appli

C 28	26.2	11.2	11707	4	US-09-136-574A-1 Sequence 1, Appli
C 29	26	11.2	1294	3	US-08-464-523B-3 Sequence 3, Appli
C 30	26	11.2	1553	3	US-08-492-459-21 Sequence 21, Appli
C 31	26	11.2	1553	3	US-08-423-752-21 Sequence 21, Appli
C 32	26	11.2	1553	4	US-08-716-873-35 Sequence 35, Appli
C 33	26	11.2	1553	4	US-09-368-431-35 Sequence 35, Appli
C 34	26	11.2	1553	4	US-09-414-006-21 Sequence 21, Appli
C 35	26	11.2	1776	1	US-08-464-523B-4 Sequence 4, Appli
C 36	26	11.2	2274	3	US-08-492-459-13 Sequence 13, Appli
C 37	26	11.2	2274	4	US-08-423-752-13 Sequence 13, Appli
C 38	26	11.2	2274	4	US-08-716-873-27 Sequence 27, Appli
C 39	26	11.2	2274	4	US-09-368-431-27 Sequence 27, Appli
C 40	26	11.2	2274	4	US-09-414-006-13 Sequence 13, Appli
C 41	25.8	11.1	607	4	US-09-328-111-205 Sequence 205, App
42	25.8	11.1	2733	1	US-08-676-967-3 Sequence 3, Appli
43	25.8	11.1	2733	1	US-08-676-974-3 Sequence 3, Appli
44	25.8	11.1	2733	2	US-09-098-487-3 Sequence 3, Appli
C 45	25.8	11.1	84495	4	US-09-797-906-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-619-554-1/C
; Sequence 1, Application US/08619554
; Patent No. 5821353
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS, Cameron M.
; APPLICANT: CHREBET, Gary L.
; APPLICANT: CLEMAS, Joseph
; APPLICANT: EL-SHERBINI, Mohammed
; APPLICANT: FOOR, Forrest
; APPLICANT: KAHN, Jennifer
; APPLICANT: KELLY, Rosemarie, - PARENT, S.A.
; APPLICANT: MARRINAN, Jean, - RAMDAN, N.M.
; APPLICANT: MORIN, Nancy, - REGISTER, E.A.
; APPLICANT: ONISHI, Janet, - SHEL, Gan-du
; TITLE OF INVENTION: DNA ENCODING 1,3 BETA-D GLUCAN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JOSEPH A. COPPOLA - MERCK & CO., INC.
; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA: US/08/619, 554
; APPLICATION NUMBER: US/08/619, 554
; FILING DATE: 01-AUG-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: COPPOLA, JOSEPH A
; REGISTRATION NUMBER: 38, 413
; REFERENCE/DOCKET NUMBER: 19104PI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-6734
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7655 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-619-554-1

Query Match 12.8%; Score 29.8; DB 1; Length 7655;
Best Local Similarity 54.0%; Pred. No. 1.5;
Matches 61; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 13 agtgaataatgtagtgctcctctgtatgagtggaagatttgattatgaatgagc 72
Db 4310 AGCGAATAGATTCTTGCTCTCACTTCAGTCAAGGCGCTTCATCCAGTAGGCG 4251

QY 73 aatctctgagatgagctcagaacaaatgactcggaagctacccagat 125
Db 4250 AATTGTGAAGTCTGGTAAGCTCTCAGCAAAACTCAGCATTTTCCAGTTTCA 4198

RESULT 2
US-09-289-254-1/c
Sequence 1, Application US/09289254
Patent No. 6299882
GENERAL INFORMATION:
APPLICANT: Junker, David E.
TITLE OF INVENTION: No. 6299882el Recombinant and Mutant Herpesviruses
FILE REFERENCE: SI0994
CURRENT APPLICATION NUMBER: US/09/289,254
CURRENT FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 2596
TYPE: DNA
ORGANISM: Marek's Disease Virus 1
US-09-289-254-1

Query Match 12.2%; Score 28.4; DB 4; Length 2596;
Best Local Similarity 58.1%; Pred. No. 2.8;
Matches 50; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 90 ctctcagaanaaactgactcggaagctacccagattattttgtgacctccat 149
Db 376 CTTGTGAAGTCGCTGACATCGAAATTAAGTGTGATTTGTCGATTCCTGAC 317

QY 150 cagctcacccatcagatgagatg 175
Db 316 CGTCCGCCGCTTATGTTGGAAATG 291

RESULT 3
US-08-976-259-70/c
Sequence 70, Application US/08976259
Patent No. 6316609
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Choi, GIL H.
APPLICANT: Welch, Rodney A.
TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli
Patent No. 6316609
NUMBER OF SEQUENCES: 142
CORRESPONDENCE ADDRESS:
ADDRESSER: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Ave, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976,259
FILING DATE: Herewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,626 AND US 60/061,953
ATTORNEY/AGENT INFORMATION:
NAME: Scieffe, Eric K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488, 0740002/EKS/CBM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 17710 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-976-259-70

Query Match 12.2%; Score 28.4; DB 4; Length 17710;
Best Local Similarity 50.7%; Pred. No. 6.6;
Matches 68; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 100 catgactcgggaactacccagattattttatgcatgacctccatcagctacc 159
Db 12783 CGAAGACGATMACCTGCGGCGCATTTTGTGATGGTGGATTCAGCCGCCCTCC 12724

QY 160 atcagatcgagatgataatgcatatgcatgagctcagctcattcacaacacggcttgagc 219
Db 12723 GACCATCTGACCGGTGACACATTTTCAGACTCCGACCACTTCTGACCGGTTGAC 12664

QY 220 aagggttatgtgtg 233
Db 12663 GATGTTCCGTTG 12650

RESULT 4
US-09-345-217-3
Sequence 3, Application US/09345217
Patent No. 6268142
GENERAL INFORMATION:
APPLICANT: DUFF, GORDON W.
APPLICANT: COX, ANGELA
APPLICANT: CAMP, NICOLA J.
APPLICANT: DIGIOVINE, FRANCESCO S.
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS FOR DISEASES ASSOCIATED
WITH AN IL-1 INFLAMMATORY HAPLOTYPE
FILE REFERENCE: MSA-010, 02
CURRENT APPLICATION NUMBER: US/09/345,217
EARLIER FILING DATE: 1999-06-30
EARLIER APPLICATION NUMBER: PCT/GB98/01481
EARLIER FILING DATE: 1998-05-21
EARLIER APPLICATION NUMBER: 9711040.7
NUMBER OF SEQ ID NOS: 32
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 3
LENGTH: 12565
TYPE: DNA
ORGANISM: Homo sapiens
US-09-345-217-3

Query Match 12.1%; Score 28.2; DB 4; Length 12565;
Best Local Similarity 53.1%; Pred. No. 6.6;
Matches 60; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 85 tgaagactcagaanaaactgactcggaagctacccagattattttatgcatgacc 144
Db 12129 tgcctcctgacatgtagagctctcgcacttggaacttgatatgaagatgctgtgc 12188


```

? APPLICATION NUMBER: US/08/936,165A
? FILING DATE: 24-SEP-1997
? CLASSIFICATION: 536
? PRIOR APPLICATION DATE:
? APPLICATION NUMBER: 60/027,032
? FILING DATE: 24-SEP-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Glanvil, Edward R
? REGISTRATION NUMBER: 38, 891
? REFERENCE/DOCKET NUMBER: P50549
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 610-270-4478
? TELEFAX: 610-270-5090
? TELEX:
? INFORMATION FOR SEQ ID NO: 8:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1172 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: Genomic DNA
? US-08-936-165A-8

```

	Query Match	Best Local Similarity	51.8%;	Score 27.4;	DB 4;	Length 1172;
	Matches	52;	Conservative	0;	Mismatches	41;
					Indels	0;
					Gaps	0;
Qy	49	aggatttgagttatgaaagagacatctcagatagagacatcagaacaatgac	108			
Db	605	AGCAATGGCTATTTCATTAATAAAGAGTATTGTATTAACCTTGACAAAGTAACGGGT	664			
Qy	109	gggaagctacgcgacgattatttatttgcatt	141			
Db	665	GATCCAGTCCCAAGAGTACTTAATTAACAAGTT	697			

```

RESULT 9
US-09-040-485-1
; Sequence 1, Application US/09040485
; Patent No. 6166176
; GENERAL INFORMATION:
; APPLICANT: Radosevlch, James A.
; TITLE OF INVENTION: A GENE ENCODING A NOVEL MARKER FOR
; TITLE OF INVENTION: CANCER
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BRINKS, HOFER, GILSON & LIGONE
; STREET: NBC Tower - Suite 3600, 455 N. Cityfront
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60611-5599
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,485
; FILING DATE: 17-MAR-1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Martin, Alice O.
; REGISTRATION NUMBER: 35,601
; REFERENCE/DOCKET NUMBER: 8998/3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-321-4200
; TELEFAX: 312-321-4299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2442 base pairs

```


FILED DATE: 17 JAN 1995
ATTORNEY/AGENT INFORMATION:
NAME: ARLES A. TAYLOR, JR.
REGISTRATION NUMBER: 39,395
REFERENCE/DOCKET NUMBER: 1242/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 493-8000
TELEFAX: (919) 419-0383
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 11283
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: no
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: Homo sapiens sapiens
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE: adult
TISSUE TYPE: female breast
CELL TYPE: normal and cancerous breast cells
CELL LINE: MCF-7
ORGANELLE: no
IMMEDIATE SOURCE:
LIBRARY: cDNA library derived from human
CLONE: obtained using published sequence
POSITION IN GENOME:
CHROMOSOME/SEGMENT: unknown
MAP POSITION: unknown
UNITS: unknown
FEATURE:
NAME/KEY: BRCA2
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: gene encoding BRCA2 protein
PUBLICATION INFORMATION:
AUTHORS: Wooster, R. et al.
TITLE: Identification of the breast cancer
JOURNAL: Nature
VOLUME: 379
PAGES: 789-792
DATE: 1995
RELEVANT RESIDUES IN SEQ ID NO: 3
US-09-099-753-3

Query Match 11.8%; Score 27.4; DB 3; Length 11283;
Best Local Similarity 65.6%; Pred. No. 12;
Matches 40; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 48 aagagatttgatgtatgaagagcaatctctgagtatgagacttaagaacaatgact 107
DB 6579 AAGAAATTTAATTAATCAATAAATTAAGTGTGAAGGTGTTCTTCAGAAATAATCACT 6638
QY 108 c 108
DB 6639 C 6639

RESULT 14
US-08-986-106-3
Sequence 3, Application US/08986106
Patent No. 6177410
GENERAL INFORMATION:
APPLICANT: HOLT, JEFFREY T.
APPLICANT: JENSEN, ROY A.
APPLICANT: KING, MARY-CLAIRE
APPLICANT: STEINER, MITCHELL S.
APPLICANT: ROBINSON-BENION, CHERYL L.

APPLICANT: THOMPSON, MARILYN E.
TITLE OF INVENTION: THERAPEUTIC METHODS FOR
PROSTATE CANCER
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARLES A. TAYLOR, JR.
STREET: SUITE 1401, UNIVERSITY TOWER, 3100 TOWER
STREET: BOULEVARD
CITY: DURHAM
STATE: NORTH CAROLINA
COUNTRY: USA
ZIP: 27707
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage
COMPUTER: IBM PC/XT/AT compatible
OPERATING SYSTEM: Windows 3.1
SOFTWARE: WORD PERFECT 6.1 and ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/986,106
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/603,753
FILING DATE: 20 FEB 1996
ATTORNEY/AGENT INFORMATION:
NAME: ARLES A. TAYLOR, JR.
REGISTRATION NUMBER: 39,395
REFERENCE/DOCKET NUMBER: 1242/3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 493-8000
TELEFAX: (919) 419-0383
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 11283
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: BRCA2
PUBLICATION INFORMATION:
AUTHORS: Wooster, R. et al.
TITLE: Identification of the breast cancer
JOURNAL: Nature
VOLUME: 379
PAGES: 789-792
DATE: 1995
US-08-986-106-3

Query Match 11.8%; Score 27.4; DB 4; Length 11283;
Best Local Similarity 65.6%; Pred. No. 12;
Matches 40; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 48 aagagatttgatgtatgaagagcaatctctgagtatgagacttaagaacaatgact 107
DB 6579 AAGAAATTTAATTAATCAATAAATTAAGTGTGAAGGTGTTCTTCAGAAATAATCACT 6638
QY 108 c 108
DB 6639 C 6639

RESULT 15
US-08-639-501-1
Sequence 1, Application US/08639501
Patent No. 5837492
GENERAL INFORMATION:
APPLICANT: Tavligian, Sean V.
APPLICANT: Kamb, Alexander
APPLICANT: Simard, Jacques
APPLICANT: Couch, Fergus
APPLICANT: Kommens, Johanna


```

; APPLICANT: Weber, Barbara
; TITLE OF INVENTION: Chromosome 13-Linked Breast Cancer
; TITLE OF INVENTION: Susceptibility Gene
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti
; STREET: 1201 New York Avenue N.W., Suite 1001
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 22204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/639,501
; FILING DATE: 29-APR-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/585,391
; FILING DATE: 11-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/576,559
; FILING DATE: 21-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/575,359
; FILING DATE: 20-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/573,779
; FILING DATE: 18-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24884-116802-04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 229..10482
;
US-08-639-501-1

Query Match      11.8%; Score 27.4; DB 2; Length 11385;
Best Local Similarity 65.6%; Pred.No.12;
Matches 40; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 48 aaggaattgagtatgaatgaagcaatctctgagatgagacttaagaagaacatgact 107
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 6611 AAGCAATTAAATATCAATAACTTAATGTTCAGAGTGCTGTCAGAAATAATCACT 6670
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 108 c 108
DB 6671 c 6671

Search completed: July 3, 2002, 07:40:58
Job*Time: 24023 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 05:07:00 ; Search time 5254.53 Seconds

(without alignments)
1047.416 Million cell updates/sec

Title: US-09-300-482-4

Perfect score: 263
Sequence: 1 gaaagacatttgatgtgtgc.....agngnactnnnnnganna 263

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapept 1.0

Searched: 1797656 segs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_in:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match Length	ID	Description
TITLE				

1	126.4	48.1	2161	8	AY056232	AY056232 Arabidops
2	76.2	29.0	95865	8	AC003970	AC003970 Arabidops
3	68.6	26.1	2058	8	STG6PDHPI	X63923 S. tuberosum
4	66.6	25.3	1396	8	SO000184	AJ000184 Spinacia
5	66.6	25.3	2049	8	SO000182	AJ000182 Spinacia
6	65.2	24.8	580	8	AF260736	AF260736 Cucurbita
7	62.2	23.7	1869	8	NTT6C18	AJ001772 Nicotiana
8	61.4	23.3	10195	1	U32737	U32737 Haemophilus
9	61	23.2	1630	8	ATG6PDHES	X84229 A. thaliana
10	61	23.2	2114	8	AY065042	AY065042 Arabidops
11	60.6	23.0	1942	8	NTT6C16	AJ001771 Nicotiana
12	60	22.8	1975	8	ATJ00135	AJ001359 Arabidops
13	59	22.4	1772	8	ATG6PDH15	X84230 A. thaliana
14	58.2	22.1	1976	1	AL591981	AL591981 Listeria
15	56.8	21.6	1976	8	CCA6246	AJ006246 Cyanidium
16	56.6	21.5	463	3	ASU09032	U09032 Anastrepha
17	56.4	21.4	11477	1	AE006191	AE006191 Pasteurel
18	54	20.5	3015	6	AF269359	AF269359 Staphyloc
19	54	20.5	3015	6	AX144679	AX144679 Sequence
20	54	20.5	3081	1	AF269560	AF269560 Staphyloc
21	54	20.5	3081	6	AX144880	AX144880 Sequence
22	53.6	20.4	1670	3	AB021914	AB021914 Anastreph
23	53.4	20.3	313450	1	AL596170	AL596170 Listeria
24	53	20.2	4942	1	AF417113	AF417113 Enterococ
25	52	19.8	35000	8	SPAC9	AL121764 S. pombe
26	52	19.8	37013	8	SPAC3A12	Z55395 S. pombe chr
27	51.8	19.7	2049	3	S67872	S67872 Zw-glycose-
28	51.8	19.7	135551	1	D90906	D90906 Synechocyst
29	51.4	19.5	1808	1	CTU83195	U83195 Chlamydia t
30	51.4	19.5	12333	1	AE002192	AE002192 Chlamydia
31	51.4	19.5	16207	1	AE002314	AE002314 Chlamydia
32	51.4	19.5	27650	1	AC090967	AC090967 Staphyloc
33	51.4	19.5	301550	1	AP003134	AP003134 Staphyloc
34	51.4	19.5	346900	1	AP003362	AP003362 Staphyloc
35	51	19.4	2750	3	PFAC6PD	M80655 Plasmodium
36	51	19.4	3244	3	PFAC6PD	X74988 P. falciparu
37	50.6	19.2	2055	4	MRU13899	U13899 Macropus ro
38	50.6	19.2	2359	8	AF012861	AF012861 Petroseli
39	50.6	19.2	2494	6	AR149300	AR149300 Sequence
40	50.2	19.1	1810	8	AY065054	AY065054 Arabidops
41	50.2	19.1	1862	8	ATH010970	AJ010970 Arabidops
42	50.2	19.1	10548	1	AE004417	AE004417 Vldrio ch
43	49.8	18.9	1690	8	ATH010971	AJ010971 Arabidops
44	49.8	18.9	1953	8	NTG6PD	X99405 N. tabacum m
45	49.8	18.9	2098	8	STU010712	AJ010712 Solanum t

ALIGNMENTS

RESULT 1	AY056232	2161 bp	mrna	linear	PLN 30-SEP-2001
LOCUS	Arabidopsis thaliana putative Glucose-6-phosphate dehydrogenase				
DEFINITION	(F14J9.8) mRNA, complete cds.				
ACCESSION	AY056232.1	GI:15810386			
VERSION	AY056232.1	GI:15810386			
KEYWORDS	FLI CDNA.				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.				
AUTHORS	1 (bases 1 to 2161) Yamada,K., Banh,J., Banno,F., Dale,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinozaki,K., Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.				
TITLE	Full Length cDNA of gene F14J9.8 (GI:3482917)				

JOURNAL 2 Unpublished
REFERENCE 2 (bases 1 to 2161)
AUTHORS Yamada, K., Banb, J., Banno, F., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koeseema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Tracy, S.E., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
TITLE Direct Submission
JOURNAL Submitted (12-SEP-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Banb, J., Banno, F., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koeseema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tracy, S.E., Davis, R.W., Ecker, J.R. and Theologis, A.
Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.
FEATURES
source
1. 2161
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL08-11-122 (R11098)"
/note="This clone is in a modified pBluescript vector (FIC-1) as a BamHI/XhoI insert."
ecotype: Columbia
1. 2161
/gene="F14J9.8"
1. 89
/gene="F14J9.8"
90. 1967
/gene="F14J9.8"
/codon_start=1
/evidence=experimental
/product="putative Glucose-6-phosphate dehydrogenase"
/protein_id="NAL07081.1"
/db_xref="GI:15810387"
/translation="MSISGLIPRSGSATAPSSVCSCHTAAASNFPVSSRDYSRSGSLVLYGGNSNLCRRGCKLMTLKLNRGNNKHQVNEITLTHSKTPTLSDDERGAEETAEEDRREPENTLGLDNGDFHNVGDLPVNSQLSDSLDVSRRASLCAVAGATGELARKITPPLFALYISGYLPEDVAITGVSKNLTJDDLSLIIASLTLCVADHACGGKMDAFOFSRTYYINGVYNNRDMGRMLERKMKOLEGSEANRIFLVSPOALVDYACIGDNAQAQPRGWTIRIVKPEGFNSHSHQTLKSLSEKFEQIYRIDHMGRLNILENTLRFNSLVEFENRRTYIRNIOVISSEISIAQPEKESDGIYRIDVSHSLIQTALLAMEPRISLDGEDINEKVKYIRKRIIDPDVILGCKSSSRKDKNGITLNGVDPTYCAAATYIDNARDGVEFLYRVGTGLKHVEVILHVCFRHVGMLVYRNETINIDLTGNELIRDPEDAILIKVINKVYFGJGLDQASLMLKDRYKTEPVSYEHLIDHVDIGDNHLFMKSDVAAAMNLTLPVLEIDKHNHTAPBLYERGGRPVAAYYLAKHGVPWAD D"
1968. 2161
/gene="F14J9.8"
misc_difference 2144
/gene="F14J9.8"
/note="compared to genomic sequence"
/replace="t"
BASE COUNT 651 a 429 c 492 g 589 t
ORIGIN

Query Match 48.1%; Score 126.4; DB 8; Length 2161;
Best Local Similarity 66.8%; Pred. No. 1.9e-23;
Matches 169; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 1 gaagacatttggagctggtgcacatgcttctggaacagatgctcagaccacaaggatgg 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 900 GAAGCTGTGCTGAGTGTGCATGCACCATGGAGATACGCTCAGCCACGAGGCTGG 959
QY 61 aatgcataataattgagaagacattgcttgatgcacttctcccataggctgaca 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 960 ACTCGAATATATGTTGAGAAACCTTTGTGTTTACCTACATTCGCTCATCATGTAA 1019
QY 121 caatcttcttcaacttccaggaagaacaaatataatagaaatgancatctactaga 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1020 AAGTACTTCTCTCTTAAGTTGAAAGAAATCTACAGATAGATACATGTGTA 1079
QY 181 aggaatcagaaatnaaaacccncaagtttaagggttcaannnagtttgaagccact 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1080 AGAAACCTCATGGAATCTGACACGTGTAAAGTTTCAATCTATTGTAACCACTA 1139
QY 241 tngagngnacct 253
Db 1140 TGAACACGACAT 1152
RESULT 2
AC003970/c
LOCUS
DEFINITION Arabidopsis thaliana chromosome I BAC F14J9 genomic sequence
ACCESSION AC003970
VERSION AC003970.1 GI:3176695
KEYWORDS HTG.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 95865)
AUTHORS Au, M., Araujo, R., Buehler, E., Dewar, K., Feng, J., Kim, C., Li, Y., Oji, O., Osborne, B. I., Shinn, P., Sun, H., Toriumi, M., Wyotskaka, V., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
JOURNAL Submitted (31-DEC-1997) Biochemistry, Stanford University/DNA Sequencing and Technology Center, 855 California Avenue, Palo Alto, CA 94304, USA
TITLE Direct Submission
JOURNAL Submitted (03-JUN-1998) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE 4 (bases 1 to 95865)
AUTHORS Federspiel, N.A., Palm, C.J., Conway, A.B., Kurtz, D.B., Conway, A.R., Au, M., Araujo, R., Buehler, E., Dewar, K., Feng, J., Kim, C., Li, Y., Oji, O., Osborne, B. I., Shinn, P., Sun, H., Toriumi, M., Wyotskaka, V., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
TITLE Direct Submission
JOURNAL Submitted (28-AUG-1998) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE 5 (bases 1 to 95865)
AUTHORS Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,

TITLE
JOURNAL

Altafi, H., Nguyen, M., Lam, B., Southwick, A., Ecker, J., Theologis, A. and Davis, R.M.
Direct Submission
Submitted (26-JUL-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

COMMENT

On Jun 3, 1998 this sequence version replaced gi:2829860.
Bases 94718-95865 of clone F14J9 overlap with bases 1-1148 of 'IGF'
clone F21M12, gb|AC000132
e-mail for correspondence: arabesequence.stanford.edu
Genes with similarity to proteins in the databases are described as
'putative', '-like' or 'similar to'. Genes that have EST
similarity but no significant protein similarity are described as
'unknown proteins'. Genes that are annotated based only on gene
prediction software are described as 'hypothetical proteins'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://combio.ornl.gov/section/index.html), GENSCAN (Chris Burge,
http://genomic.stanford.edu/~chris/GENSCANW.html), Fexa (V.Solovayev
& A.Salamov, Sanger Centre, http://genomic.sanger.ac.uk/), and
NetPlantene (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/NetPlantene.html).

FEATURES
source

1. 95865
/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="1"

/clone="F14J9"

/note="contains phyA molecular marker"

/gene="F14J9.1"

join(1491..1796,1887..2210,2364..2515,2595..2817)

/gene="F14J9.1"

/note="Similar to rice water stress induced protein
g11537404"

/codon_start=1

/protein_id="AAC33195.1"

/db_xref="GI:3482910"

/translation="MAPEMNNKLSYGEKKRAYVFLTAGDGYKVGVLAKGLKRTKS
KPLVAVLPDVPADHRRLDGGVKEIOVYPDPNOQOFAMAYVLANSKRIK
EVEYSKRLIYDGOVPEINDHLEFDLPDGNFYANKDCCEKTSHTPOYKIGYQCCP
DKYTPPESEIGPPRPILYFNAGMYEPESLDPTINLETLKVPPTPRAEODEFLMYIK
DIKKPIPPVNLVLAMLMRHPEINLEAKVVAHYCAGAKPWRPTGOEENFERDIDM
LVERKMDIYNDESIDYKFNHCGQKEDVHRKPTLPQPFITDSEADVLQAKAPSSA
"

3405..3971

/gene="F14J9.2"

/note="Hypothetical protein"

/codon_start=1

/protein_id="AAC33196.1"

/db_xref="GI:3482911"

/translation="MIVLVLYITFTPHIEGVADGPTEALINSICIENEDYGFCKIIT
EKKTPTATIKETLSLIYTTMAASDPIYFINDILRKWGPKEKTSGLKCHAYKKE
TNSFLIRFLFSKGEYELMDPAISTAKILEECRGDPLIPYKLEKKRVMRLIT
MSAVSGHMKNEKASLISVVAQFNI"

6504..7028

/gene="F14J9.3"

/note="Hypothetical protein"

/codon_start=1

/protein_id="AAC33197.1"

/db_xref="GI:3482912"

/translation="MHHLISIVLAIIASSVAEAFIDEKTOQRYNGICKQMDTRFCS
SLIKINLTFPASKETIMNTVSEAEERFANTEFFISTLLRNAGDERPDIOACAEAA
IVNSAFKAVTFEFOAYYSKIYINIEKRVSMVADICKDFNVLVGQINPLLEKNQGTI
LISEOIVHWMS"

complement(7954..10925)

/gene="F14J9.4"

complement(join(7954..8155,8251..8402,8490..8648,

8762..9014,9101..9217,9294..9356,9577..9606,9895..9938,
10404..10444,10829..10925))

/gene="F14J9.4"

/note="Similar to Mtn21, g112598575, Megicaco truncatula
nodulation induced gene"

/codon_start=1

/protein_id="AAC33198.1"

/db_xref="GI:3482913"

/translation="MAKSDMLPFLAMVVOIGVAGNITSKMAEANDETIFINIV
KMYKKNVKKKTKKRCRAIFICODITRKRITLRLIVGVFCSTIGANGQVLYEV
GLONSPPIACALNPLPAVFLIARVIOETVAKKASGQAKYIGLVGIVGAMYS
FYHHTTGIGSKTHMVAENITVTHGSSGNFLEPPLIIMAAVMAAMEIOTKM
SETPAAPYTSLLKMGSLDCCGALISDITSDMSLSPLRISLNYAGVASALA
FCLSMAMQRRGPILVVSFPLILVAIVTSWALEKLTGTFTGSAVLYGLTGLV
WGMKREVSKEEREKVKQOHNKKSSEMEDIESRLPVASSGNGSTRSP"

12867..15018

/gene="F14J9.5"

join(12867..13122,13720..13911,13990..14156,14269..14533,
14786..15018)

/gene="F14J9.5"

/note="Similar to nodulins and lipase; location of EST
E6C297, gb|AA042309.. similar to nodulins g113328240,
g112129854 and others and lipase, g112129636"

/codon_start=1

/protein_id="AAC33199.1"

/db_xref="GI:3482914"

/translation="MATLSLSHSHSFLVLLPFLILRONTAVAGCCOVPIYFNFGDS
NSDVGVLAVAGISIGLPNGRSPFORSTGRSDRLIDFICOLNLSLNPYDLSV
GSKRONGANFAIVGSSITLPYRFPALNLOMFHFSRALEIASIDPKEMIGDS
GFRNALYMIIDGONDIDSPSKGYSRYVKKLIPNVSEIKSAKITIYDEGRFPWY
NTPGLCPQKQKSNVSHSGFEDKHCGLTNNAAALFNEGDLHMGCRDITREKENTY
VDIYAIRYDILANSNNYGFEPKMACCGYGPPYNNVNTCGNGSKSCDEGSRFIS
WDGIHYETANAIVAMKVLVSMQSHSTPPTPFHFCG"

complement(15279..16596)

/gene="F14J9.6"

complement(join(15279..15796,15879..16037,16122..16252,
16328..16468,16571..16596))

/gene="F14J9.6"

/note="Similar to 12-oxophytodienoate reductase,
g112765083 and old-yellow-enzyme homolog, g112232254"

/codon_start=1

/protein_id="AAC33200.1"

/db_xref="GI:3482915"

/translation="MKNFNLTNRITVMAPMARMSYGNIPQPHVALLYCQRTPGGLLI
SEATGSEETAMAYNMGIMRKEQIEAMKPLVDVAHSHGGLFECQLMHAGVSDQDC
PNGESPVSTDKPPADDPNSNEFPTRLRDEIPTIINDPLARNMTEAGEDEVEH
GARGYLIDQEPKDSVNDRTDSYSGSLNRCGRFALOVLEAVSKETGPRVGRISPPAD
YMESGDTDPKRLGLYMAKSLNRFETILCHMIEPRMKYVSEIFECRESLTPRRNAFNGT
FIVAGGYTREDGNKNAVAGRTDVAVGRFLPLANDLPKRELPLNPK"

17046..19163

/gene="F14J9.7"

17046..19163

/gene="F14J9.7"

/note="Hypothetical protein: Similar to Arabidopsis
selenium-binding protein, g112244760 and Arabidopsis
hypothetical proteins g113033399, g112464864, g112244839
and several others"

/codon_start=1

/protein_id="AAC33201.1"

/db_xref="GI:3482916"

/translation="MKSQILRLRTYSTTIPPTANVRITHLISRIKTHEARKLFDS
SKSISNNSMVAAGFYANLIMPRDARKLEDENMDYIISMNGVSGMNGLEDEARKYF
DLMEKRVVSWTALVKGYNHGVKDVASELEKMKPEKKNKSWTMYLGLFODGDIDA
KLIYEMIPEKDNARTSMITGICEGVADARETFEDMSRSVITTTMTTGYGCONR
VDDARKTFDVPKETEVSMTLMGYVQNGRIDEAFELFEMPKYPIYACNAMISGIG
QKGEIARARVPDSMERNDASWQTVIKHERNCFELEALDPLILMOKQGRPTFPL
ISLISVCASLASLHGRQVHAQLVRCQPDVYVAVSLMTMYIKCGELVSKLIFDR
PSIKDIIWNNSTISGYSAGHGEALVKFCEPGLSGSTKPREVEFVATLSACVAGME
EGKRTYSMESVGEVCPDITAHYACVMDLGRAGFENPAMEIDSMVEPDAVWGSLE
GACRTHSOLDVAEFCARPKLIEIPENSGYTIILSNMYASOGRAVDVLEIKMTKTY
RKSFGCSWTEYENKVIHFTFGGINSHPQESILKIDLEDBGLREAGYNPDSCIALND
VDEEKVNSLTKHSERLVAVALTKISEGIFIRVMKNLRVCSDCHTAIKITSKERE
IILDRANRHFHFRNGECSCKDYW"

RESULT	3	
LOCUS	STG6PDHPI	
DEFINITION	STG6PDHPI	2058 bp mRNA linear
ACCESSION	X83392.1	PLN 15-FEB-1996
VERSION	X83392.1	
KEYWORDS	66pdh gene; glucose-6-phosphate dehydrogenase.	
SOURCE	Potato.	

FEATURES	
source	Location/Qualifiers
	1. .2058
	/organism="Solanum tuberosum"
	/cultivar="Desiree"
	/db_xref="taxon:4113"
	/clone="BBSK-platfuld #3, #4"
	/haplo_type="diploid (2n)"
	/cell_type="mesophyll"
	/tissue_type="green leaves"
	/clone_lib="lambda Zap1"
	/dev_stage="vegetative stage"
CDS	41. .1774

```

/product="glucose-6-phosphate dehydrogenase"
/protein_id="CA5875.1"
/db_xref="GI:1197385"
/db_xref="SWISS-PROT:Q43839"
/translation="MCGVRLRNPCCSSAATSPSTFHNGTPYCKENFLPERTQPLN
WVGSIQIPKHKHEEVSSNEFLPSSAVOVOVQVLTPLGSGDPTVSTVIGASDLN
AKKILPALFALPYEDCLDENFVNFVGYSTKSLDSLELRNLTITTLTKRDKRNCNAK
MHPLEFVGHSGQVNSDEDFAELDKLEKGCGRNSMLFTSLPPIPNFVYVVRAS

```

Query March	26.1%	Score 68.6;	DB 8;	Length 2058;
Best Local Similarity	52.8%	Pred. No. 4.2e-06;		
Matches 131; Conservative	0;	Mismatches 117;	Indels 0;	Gaps 0;

RESULT	4
LOCUS	SC0000184
DEFINITION	SO0000184 1396 bp mRNA linear PLN 23-JUL-1997
ACCESSION	Spinalcia oleracea mRNA for glucose-6-phosphate dehydrogenase, partial, clone O28FA38.
VERSION	AJ000184
KEYWORDS	AJ000184.1 GI:2276347
SOURCE	G6PD gene; glucose-6-phosphate dehydrogenase, spinach.

```

FEATURES      location/Qualifiers
source        1. .1396

```

```

/db_xref="taxon:3562"
/clone="O28FA38"
/tissue_type="leaves"
/dev_stage="vegetative"
1..1396
/gene="G6PD"
<1..>1396
/gene="G6PD"

```



```

/EC_number="1.1.1.49"
/codon_start=-1
/product="Glucose-6-phosphate dehydrogenase"
/protein_id="CAA03941.1"
/db_xref="GI:2276348"
/db_xref="SPTREMBL:O24359"
/translation="EFSNGHPNDVSLQNDVAVNPVAKSIDPSADLQLPLLESYKE
EPTLSIIIVGASGLDAKRLIFPLFLFYENCLEPENTVYFGSPRTKEMNDELRTMISK
TLTRCIDORENCEKMDHFLQRCFYHSGQYNSDESGLDCKLKEKAGRLQRLFYL
SIPNIEFDVYRCVSHRASASGWTIRYIVKPGRSDSSRELTRSFKOYLSDDQIFR
IDHYLKELEVENSVLRFSNLVPEPLMSRNTYNNOLIFSEDCRTEGRCGSPNNYGI
RDIMONHLLDITLAFAMETPVSLDIEDIRNEKVKVLSMKPLKLDVVGQYTGHSKG
MKSISGYTDPIVNNNSVTPTEFAAALFIDNARMWDVGPLMKAGKALHTRAEIRVOF
RHVPGNLVYKKECTGTDKATNELVLRYOPDEALYIKNNKVPGLGMRDLDRDILCY
TRYRGEIPDAYERILL"
misc_feature
959..979
/gene="G6PD"
/function="G6PD catalytic site"
BASE COUNT      418 a      252 c      329 g      397 t
ORIGIN

Query Match      25.3%; Score 66.6; DB 8; Length 1396;
Best Local Similarity 52.4%; Pred. No. 1.4e-07;
Matches 129; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 8 ttgtgagtgtgcgtcatgtcttgcagaagcgtctcagaccagaaggaatgcga 67
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 500 TTGTGAGATGTGTGAGATGTGTAGCCATCGAGCTTCTTCGCAAGTGCATGCAAGG 559

QY 68 taattattggaagcatttgcattgatgcacttctccatagagtgacacaatac 127
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 560 TTATAGTAGAAGAACCTTTTGGACGTACGATTCCTTCTCGTAGCTACTAGAAAGCT 619

QY 128 ttcttcaacttcaggaagaacaatatataagattgancatctactagaagaatc 187
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 620 TCAAGCACTACTTGAAGCAGACCATATTCAGATTGACCATTAATCTTGCGGAAGAGT 679

QY 188 ncagtnaaatccnccaggttcaagggttccaannagnnttgcaccacttngagn 247
    '|||' ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 680 TGGTGAGAAATTATTCGTCTTACGCTTTTCAATCTAGTTTGAACCTTATGCTCAA 739

QY 248 gnaact 253
    |||
Db 740 GGAAGT 745

RESULT 5
LOCUS      S0000182      2049 bp      mRNA      linear      PLN 23-JUL-1997
DEFINITION Spinacia oleracea mRNA for glucose-6-phosphate dehydrogenase,
            clones O28FA14 & O30A4.
ACCESSION  AJ000182
VERSION     AJ000182.1 GI:2276343
KEYWORDS   G6PD gene; glucose-6-phosphate dehydrogenase.
SOURCE      Spinach.
ORGANISM   Spinacia oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
REFERENCE   1 (bases 1 to 2049)
AUTHORS     Fink,A.
TITLE       Direct Submission
            Submitted (21-JUL-1997) Fink A., Laboratoire de Biochimie et de
            Physiologie Vegetale, Universita de Geneve, 3, Place de
            l'Universite, CH-1211 Geneve 4, SWITZERLAND
            2 (bases 1 to 2049)
REFERENCE   Fink,A., Diogon,T., Perroud,P.F., Crespi,P. and Greppin,H.
AUTHORS     Nucleotide sequences of glucose-6-phosphate dehydrogenase from
TITLE       Spinacia oleracea cloned by RT-PCR
JOURNAL     Unpublished
FEATURES    Location/Qualifiers
            source
            1..2049

```

```

/organism="Spinacia oleracea"
/db_xref="taxon:3562"
1..1659
/organism="Spinacia oleracea"
/plasmid="pZ1"
/cultivar="Matador"
/db_xref="taxon:3562"
/clone="O28FA14"
/tissue_type="leaves"
/dev_stage="vegetative"
147..2049
/gene="G6PD"
/EC_number="1.1.1.49"
/product="Glucose-6-phosphate dehydrogenase"
/codon_start=-1
/protein_id="CAA03939.1"
/db_xref="GI:2276344"
/db_xref="SWISS-PROT:O24357"
/translation="MEELVSCHHPLPLCLQSSVPPNCLTFPQDSACORCSHSEFNG
HPIDVSLQNDVAVNPVAKSIDPSADLQLPLLESYKEEPTLSIIIVGASGLDAKRL
IFPALFLFYENCLEPENTVYFGSPRTKEMNDELRTMISKTLTRCIDORENCEKMDH
LQRCFYHSGQYNSDESGLDCKLKEKAGRLQRLFYLSIPNIEFDVYRCVSHRAS
SASGWTIRYIVKPGRSDSSRELTRSFKOYLSDDQIFRIDHYLKELEVENSVLRS
NLVFEPLMSRNTYNNOLIFSEDCRTEGRCGSPNNYGIIRDIMONHLLDITLAFAMET
PVSLDAEDIRNEKVKVLSMKPLKLDVVGQYTGHSKGMSISGYTDPIVNNNSVTP
TEFAAALFIDNARMWDVGPLMKAGKALHTRAEIRVOF RHVPGNLVYKKECTGTDK
ATNELVLRYOPDEALYIKNNKVPGLGMRDLDRDILCYTRYRGEIPDAYERILLDAI
EGERELFIRSDKLDAMSLFTPLLKELEKVAPELYPSGRCPVGAHYLAANKHNVN
GDLSGEDS"
922..942
/gene="G6PD"
/function="G6PD catalytic site"
1660..2049
/organism="Spinacia oleracea"
/plasmid="pZ1"
/cultivar="Matador"
/db_xref="taxon:3562"
/clone="O30A4"
/tissue_type="leaves"
/dev_stage="vegetative"
2028..2036
/gene="G6PD"
2049
/polyA_site
/gene="G6PD"
BASE COUNT      596 a      381 c      458 g      614 t
ORIGIN

Query Match      25.3%; Score 66.6; DB 8; Length 2049;
Best Local Similarity 52.4%; Pred. No. 1.4e-07;
Matches 129; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 8 ttgtgagtgtgcgtcatgtcttgcagaagcgtctcagaccagaaggaatgcga 67
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 763 TTGTGAGATGTGTGAGATGTGTAGCCATCGAGCTTCTTCGCAAGTGCATGCAAGG 822

QY 68 taattattggaagcatttgcattgatgcacttctccatagagtgacacaatac 127
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 823 TTATAGTAGAAGAACCTTTTGGACGTACGATTCCTTCTCGTAGCTACTAGAAAGT 882

QY 128 ttcttcaacttcaggaagaacaatatataagattgancatctactagaagaatc 187
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 883 TCAAGCACTACTTGAAGCAGACCATATTCAGATTGACCATTAATCTTGCGGAAGAGT 942

QY 188 ncagtnaaatccnccaggttcaagggttccaannagnnttgcaccacttngagn 247
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 943 TGGTGAGAAATTATTCGTCTTACGCTTTTCAATCTAGTTTGAACCTTATGCTCAA 1002

QY 248 gnaact 253
    |||
Db 1003 GGAAGT 1008

```


LOCUS U32737 10195 bp DNA linear BCT 29-MAY-1998
DEFINITION Haemophilus influenzae Rd section 52 of 163 of the complete genome.
ACCESSION U32737.1
VERSION U32737.1 GI:1573536
KEYWORDS
SOURCE Haemophilus influenzae Rd.
ORGANISM Haemophilus influenzae Rd
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Haemophilus.
REFERENCE 1 (bases 1 to 10195)
AUTHORS Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A.,
Kirkness,E.F., Kerlavage,A.R., Bult,C.J., Tomb,J., Dougherty,B.A.,
Merrick,J.M., McKenney,K., Sutton,G., FitzHugh,W., Fields,C.A.,
Gooyne,J.D., Scott,J.D., Shirley,R., Liu,L.I., Glodek,A.,
Kelley,J.M., Weidman,J.F., Phillips,C.A., Spriggs,T., Hedblom,E.,
Cotton,M.D., Uterback,T., Hanna,M.C., Nguyen,D.T., Saudek,D.M.,
Brandon,R.C., Fine,L.D., Fritchman,J.L., Fuhrmann,J.L.,
Georgagen,N.S., Gnehm,C.L., McDonald,L.A., Small,K.V., Fraser,C.M.,
Smith,H.O., and Venter,J.C.
TITLE Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd
JOURNAL Science 269 (5223), 496-512 (1995)
MEDLINE 9530630
REFERENCE 2 (bases 1 to 10195)
AUTHORS Tatusov,R.L., Mshnegian,A.R., Bork,P., Brown,N.P., Hayes,W.S.,
Borodovsky,M., Rudd,K.E. and Koonin,E.V.
TITLE Metabolism and evolution of Haemophilus influenzae deduced from a
whole-genome comparison with Escherichia coli
JOURNAL Curr. Biol. 6 (3), 279-291 (1996)
MEDLINE 96398784
REFERENCE 3 (bases 1 to 10195)
AUTHORS White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-1995) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
REFERENCE 4 (bases 1 to 10195)
AUTHORS White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-1997) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
REMARK The H. influenzae sequence has been updated by R. Fleischmann. New
database matches have been assigned, product names have been
improved, and a number of frame shifts have been corrected. We
gratefully acknowledge the work of Tatusov et. al. We have
incorporated their annotation into the /notes fields of the
corresponding H. influenzae genes
5 (bases 1 to 10195)
REFERENCE White,O., Clayton,R.A., Kerlavage,A.R., Fleischmann,R.D.,
AUTHORS Peterson,J., Hickey,E., Dodson,R. and Gwinn,M.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-1998) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
REMARK The whole genome was shifted by 588 nucleotides for a new start
COMMENT On Sep 30, 1996 this sequence version replaced gi:1221227.
FEATURES
Source
Location/Qualifiers
1..10195
/organism="Haemophilus influenzae Rd"
/db_xref="taxon:71421"
90..917
/gene="HI0551"
90..917
/gene="HI0551"
/note="similar to GB:D10483 SP:P05637 GB:X04711 PID:216474
PID:40919 percent identity: 52.77; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=1
/product="diadenosine-tetraphosphatase (apah)"
/protein_id="AAC22209.1"
/db_xref="GI:1573537"
/translation="MATYFVGDLGGCYDELQLLERVDENPTODKIYIGDLVARGDK
SLECLRVKSLGNAAQTVYGNHDLALATADIKVAKPRDVRVDATFNAPDEDELHFWL
RHQPLVHNEKLNFLMSHAGISPDWDLTKASCAAEVQIILQHGDFHYLIENMYSQD

DRWSPDLGRLARHYIINATFRMFCYLDHREDFACKSPDKADPAELTPWENLDNPLY
KOPIYVGHNASLWDEPTPKGIYALDTGCVNNNMTMLRMEDKQFTQSAVKNSYDF
927..1550
/gene="HI0552"
927..1550
/gene="HI0552"
/note="hypothetical protein; identified by Genemark;
putative"
/codon_start=1
/transl_table=1
/product="H. influenzae predicted coding region HI0552"
/protein_id="AAC22216.1"
/db_xref="GI:1573538"
/translation="MLSLTASCSLFPNIPFYQFAOMKKFCEDIPAIKADYKLMWN
KATIOSVSAQGTGTFPAKHIESWTNGQVRAHFAFYKFEFNQNSAAIFSVLNRRL
RVCDMLHWRDSQINVOQYQNMIDQDFOFADFEIWRDESEYDDFRQVKVISEK
NLIRSDDEDFWICSTIEKALNDIDPLFTTHIQQLQPLYDCHQ"
complement(1676..3130)
/gene="HI0553"
complement(1676..3130)
/gene="HI0553"
/note="similar to SP:P70718 PID:1651212 percent identity:
91.53; identified by sequence similarity: putative"
/codon_start=1
/transl_table=1
/product="6-phosphogluconate dehydrogenase,
decarboxylating (gnd)"
/protein_id="AAC22210.1"
/db_xref="GI:1573539"
/translation="MSYKGDICVIGLAVMGONLLNMNDHGFYVAVYRTSKYDEPL
QGAKGNTICAYSLIEDLAALKERPKVMLVNRAGDVVDQTEALPHLEGGDITIDG
GNSVNPDTNRVRKALAEGKIFIGSVSGEGEGARHGSPMGQNMQVYKPIFOA
ISAKTEQDEPCDDWGEAGHGFVYVHNGIEYDMQICFAYQFLKEGTLSTYEMQ
AIFAKMTELDSTLIDITDILGKDSAGPLVEKILDTGQKGTGKGTINLDEG
IPLITFESVAPACVSSPKQDQVANAOMFGKTIIPVEDGRKVMLEAVAKALASKIS
YAGFPMIIRASPDQMDINGATALLMRECCIIRSPRLGIRDAVEANPMLVPLGSD
SYFGILENLSDMRKRYAKSIEVGIPIPCASATITFDGTSARLANLLQARDYF
GAHYERTKRGSEFHTWTRGNGNSTAYTDV"
complement(3576..4118)
/gene="HI0554"
complement(3576..4118)
/gene="HI0554"
/note="hypothetical protein; identified by Genemark;
putative"
/codon_start=1
/transl_table=1
/product="H. influenzae predicted coding region HI054"
/protein_id="AAC22217.1"
/db_xref="GI:1573540"
/translation="MTKRYNAGIHHRSIRLKHYNRSEGFYFTTCCKNECLPGHI
INQOMLNDLGNVYKQWENTPMFPPQVRIDEFYIMPVHLGIIIEIIOVYKGNLPL
OLRATQLPQKGTSGTISIVRRFAGVYSMARKNSEIFDVMQRYEHIHIDREKSYLG
IYEXIONNPIMEDQLYVD"
complement(4243..4482)
/gene="HI0555"
complement(4243..4482)
/gene="HI0555"
/note="similar to GB:M87483 SP:002009 PID:551879 percent
identity: 40.21; identified by sequence similarity:
putative"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAC22211.1"
/db_xref="GI:1573541"
/translation="MEILAYILTLAVTLNHTYIYLYEMFTIESKSGVRSQSVYIORLPG
SRDYFWLCHQPNFEDLFLRCLDYRRSFWSYDDE"
complement(4482..4973)
/gene="HI0556"
complement(4482..4973)
/gene="HI0556"
/note="similar to SP:P46016 PID:556607 percent identity:
35.82; identified by sequence similarity: putative"


```
/codon_start=1
/transl_table=1
/product="oxidoreductase, putative (devb)"
/protein_id="AAC22212.1"
/db_xref="GI:1573542"
/translation="WPSPDPSNNGEVOKLFDHIQIPAEINIRHROENEPHELR
FEELSAVIPNGFDMILIGMDIGHTSLSPHONFDEINAVIAKIPESQIISK
TAKLEQAKRIRYLVLTGESKADILKEIOTTPAENLPYPAKIKANGVTWYLDKAAV
RLL"
complement(5059..5181)
/gene="HI0557"
complement(5059..5181)
/gene="HI0557"
/feature="hypothetical protein; identified by Genemark;
putative"
/codon_start=1
/transl_table=1
/product="H. influenzae predicted coding region HI0557"
/protein_id="AAC22218.1"
/db_xref="GI:1573547"
/translation="MNYSFPQAQHAHVDAKIAQEFYISQLNHSRTYPPFWMDA"
complement(5264..6748)
/gene="HI0558"
complement(5264..6748)
/gene="HI0558"
/feature="similar to SP:P77809 PID:1651208 percent identity:
85.22; identified by sequence similarity: putative"
/codon_start=1
/transl_table=1
/product="glucose-6-phosphate 1-dehydrogenase (zwf)"
/protein_id="AAC22213.1"
/db_xref="GI:1573543"
/translation="MQTDNMCIVIFGASGDLTHRKLLPALYLYKIRLSNFSVLG
ARSLNDTFREKMRKALIHNETTPETLDACSHLYQAVNTSDAODYGLVRLD
LHDYQOTGNTFYNSPPLSYGLVIECLAHGLENTEYGRKRIIEKPEYDEKTD
TLDDYQIHFEFEHOYRIDHYLGKETYONLLVLRPSNGFEPILNRNFDYVETGAE
SIGVEERGQYVSGVAMDFONHLLQYLAHMEPRATITANSRDEAVAKVMSLRP
LTSDEMNNVILGQYTAETNGKMEKGLERKGPANSRTETVIALKEIENRWAG
PFYRTGKRLPARVETIVIFKTPPHVFSQNAENKILIRIODEAISMFGGLKKG
AGFEAKESMDFRYADLAGAVALYAEERLLDPAKGDATLFARTDAVAHAKFVQPL
DYKANGGRHIEYEAGTWGPVAAADKLIAKQKVMKPSGLMKKY"
complement(6824..7510)
/gene="HI0559"
complement(6824..7510)
/gene="HI0559"
/feature="similar to SP:P22255 GB:M5170 GB:M80795 PID:145275
PID:145690 percent identity: 36.07; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=1
/product="sulfolite synthesis pathway protein (cysQ)"
/protein_id="AAC22214.1"
/db_xref="GI:1573544"
/translation="MTEADLFVSQFLTEKLTALFPPNVLSEENCHISFERKKNWKEY"

Query Match 23.3%; Score 61.4; DB 1; Length 10195;
Best Local Similarity 52.1%; Pred. No. 3.2e-06;
Matches 110; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
```

```
Db 6147 AACCGCTTGAACCGTATTCATTGATTA 6117

RESULT 9
ATG6PDHE5 1630 bp mRNA linear PLN 31-JAN-1996
LOCUS A.thaliana mRNA for glucose-6-phosphate dehydrogenase (clone E5).
DEFINITION X84229
ACCESSION X84229
VERSION X84229.1 GI:1166404
KEYWORDS 96pH gene; glucose-6-phosphate dehydrogenase.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1630)
Fink,A., Greppin,H. and Tacchini,P.
Nucleotide sequence of a cDNA encoding the glucose-6-phosphate
dehydrogenase from Arabidopsis thaliana
Unpublished
2 (bases 1 to 1630)
Fink,A.F.A.
Direct Submission
Submitted (23-JAN-1995) A.F.A. Fink, Lab de Biochimie et de
Physiologie, Vigintales, Universiti de Geneve, 3 Place de
l'Universiti, 1211 Geneva, SWITZERLAND
FEATURES
Location/Qualifiers
1..1630
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/clone="E5"
/cell_line="A. thalian Co"
/tissue_type="leaves, stems, some flowers and roots"
/clone_lid="lambda Zap"
/dev_stage="adult"
<1..1630
/gene="G6PDH"
1..1630
/gene="G6PDH"
<1..1480
/gene="G6PDH"
/ec_number="1.1.1.49"
/codon_start=2
/product="glucose-6-phosphate 1-dehydrogenase"
/protein_id="CAA59011.1"
/db_xref="GI:1166405"
/db_xref="SPTREMBL:O43728"
/translation="QSOSTVSTIVGASGDLAKKIFPALFYEGCLPEHFTVFGY
SRKMTDVELRNWVSKTLTCRIDRANCGEKMEFLRCFYISGQYDSQEHFELDCK
LRKHEGRISNRLEFLYSIPNIFVDVAKCASTASSVNGRTVRYIEKPGRDETSAA
LTRKSLQYLEDQDIFRIDHYLGKELVENLSLFRSNLIEPELMSRQYIRNVQFISED
FTGREGGQYFDNNGIIRDIQONHLLQIILFAMETPVSIDAEIDREKXKVLRSMPRI
RVDVIVGQYKSKTKGQYPRATYDTPKTYPKGSLTPPAAALFIDNAWDGVPFLMK
AGKALHTRSALFIVQRRHVPNGNLINRNTGSDIDQATNELVIRQDEALYFIRNNVP
GLGMRIDRSNLINLLYSARYSKETIPDAVERLLDLAIEGERRLFRSDELDAAMSLFTPL
LKEIEEKRIPEYEPYSGRNPVGANHYLAHKVQWQDVSIDQ"
1600..1630
/gene="G6PDH"
1630
/gene="G6PDH"

polyA_signal
polyA_site
BASE COUNT 489 a 292 c 366 g 483 t
ORIGIN

Query Match 23.2%; Score 61; DB 8; Length 1630;
Best Local Similarity 52.1%; Pred. No. 4.4e-06;
Matches 124; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
```


QY	66	catatatttgggaagccatttggctttgatgcactttcttccataagctaacaca	125
	1		111
	2		111
	3		111
	4		111
	5		111
	6		111
	7		111
	8		111
	9		111
	10		111
	11		111
	12		111
	13		111
	14		111
	15		111
	16		111
	17		111
	18		111
	19		111
	20		111
	21		111
	22		111
	23		111
	24		111
	25		111
	26		111
	27		111
	28		111
	29		111
	30		111
	31		111
	32		111
	33		111
	34		111
	35		111
	36		111
	37		111
	38		111
	39		111
	40		111
	41		111
	42		111
	43		111
	44		111
	45		111
	46		111
	47		111
	48		111
	49		111
	50		111
	51		111
	52		111
	53		111
	54		111
	55		111
	56		111
	57		111
	58		111
	59		111
	60		111
	61		111
	62		111
	63		111
	64		111
	65		111
	66		111
	67		111
	68		111
	69		111
	70		111
	71		111
	72		111
	73		111
	74		111
	75		111
	76		111
	77		111
	78		111
	79		111
	80		111
	81		111
	82		111
	83		111
	84		111
	85		111
	86		111
	87		111
	88		111
	89		111
	90		111
	91		111
	92		111
	93		111
	94		111
	95		111
	96		111
	97		111
	98		111
	99		111
	100		111
	101		111
	102		111
	103		111
	104		111
	105		111
	106		111
	107		111
	108		111
	109		111
	110		111
	111		111
	112		111
	113		111
	114		111
	115		111
	116		111
	117		111
	118		111
	119		111
	120		111
	121		111
	122		111
	123		111
	124		111
	125		111
</			

5' UTR	1' UTR	607 a	399 c	452 g	656 t	3' UTR
CDS						BASE COUNT
						ORIGIN
						Query Match
						Best Local Similarity 52.1% Pred. No. 4.4e-06;
						Matches 124; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
						6 acccttgatcttcgcatatgtcttcgaagaagctcgaagccgaagaagatgaatcg 65
						Db 824 ATTGTGATGATCTGTGAAATGTGCAAGACACTTTCCTTCATCTGTCAATGATGATGACTAG 883
						66 cataatattgagaagccatttggcttgaatgacacattctccatagaatgacacata 125
						Db 884 GGTATGCTGCAGAAACCTTTGGGCGAGATTGTGAAACCTCGCGCTGCTTAACGAATC 943
						Qy 126 tctcttccaacttcgaagaaagcaatatatagaattgancatctcactagaagaa 185
						Db 944 CCTCAAGCAGTATTTGGAGGAATCAAAATTTTATGATAGACCACTTAACCTTAGGAAGA 1003
						Qy 186 tcncaagtaaaatctcnaaggttaaggtttcaaanmagnttttgaagccacttng 243
						Db 1004 GCTAGTCGAAACCTTATCTGTTCTTCGATTCOTCAACCTTATTTGAGCGGCTATGG 1061
RESULT 11						
NTTpg16						
LOCUS						
DEFINITION						
ACCESSION						
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
MEDLINE						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
REMARK						
REFERENCE						

AUTHORS vonSchaewen,A.
TITLE Direct Submission
JOURNAL Submitted (23-SEP-1998) vonSchaewen A., Plant Physiology,
University Osnabrueck, Barbarastr. 11, D-49069 Osnabrueck, GERMANY
COMMENT On Sep 25, 1998 this sequence version replaced gi:3021530.
FEATURES
Location/Qualifiers

source

1.1942
/organism="Nicotiana tabacum"
/cultivar="Samsun NN"
/db_xref="taxon:4097"
/cell_type="mesophyll"
/tissue_type="young sink leaves"
41..232
sig_peptide
41..1801
mrna
/note="TPG16"

mat_peptide
/product="glucose-6-phosphate dehydrogenase"
233..1798
/product="glucose-6-phosphate dehydrogenase"
BASE COUNT 584 a 353 c 442 g 563 t
ORIGIN

Query Match 23.0%: Score 60.6; DB 8; Length 1942;

Best Local Similarity 50.8%: Pred. NO.5.6e-06;
Matches 126; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 6 actttggaatgttcgtcgtcttgcgaagcagtcgtccagaccaggaaggaatgcg 65
DB 694 ACTTCGGATGTGTGAGATGTCAGTCTTAAGCTTCTTACACAGAGGGTGGAGCAG 753
QY 66 cataaattggaagcattggtgcttgcacatttcccatagcgtacacata 125
DB 754 AGTCATTGTGAGAAACCTTTGTCGTGATCTAGCATCTTACCGACCTACCAAGAG 813
QY 126 tcttcttcaaaccttcagaaacaaatataatagatcagatcactctactaggaagaa 185
DB 814 CCTAAGAGAGTACCTAACAGAGAGCAATATCCGAATTGACCACTTACTTGGAAAGCA 873
QY 186 tccacgttaaaatccttcacaggtttaagggttcaaanmagnitttgagccacttngag 245
DB 874 ACTTGTTGAGAACTCTCGTCTTCTTCGATCTTGTCTTGTGAGCCTCTGTGTC 933
QY 246 ngnacnt 253
DB 934 CAGAAACT 941

RESULT 12

ATJ00135

LOCUS 1975 bp mRNA linear PLN 17-AUG-1999
DEFINITION Arabidopsis thaliana cDNA encoding plastidic glucose-6-phosphate
dehydrogenase.
ACCESSION AJ001359

VERSION AJ001359.1 GI:3021304
KEYWORDS glucose-6-phosphate dehydrogenase.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 1975)
Wendt,U.K., Haenschild,R., Lange,C., Pieterama,M., Wenderoth,I. and
von Schaewen,A.
TITLE Evidence for functional convergence of redox regulation in G6PDH
isoforms of cyanobacteria and higher plants
JOURNAL Plant Mol. Biol. 40 (3), 487-494 (1999)

REFERENCE

AUTHORS

TITLE

Evidence for functional convergence of redox regulation in G6PDH
isoforms of cyanobacteria and higher plants
JOURNAL Plant Mol. Biol. 40 (3), 487-494 (1999)

REFERENCE 2 (bases 1 to 1975)
von Schaewen,A.
TITLE Direct Submission
JOURNAL Submitted (18-SEP-1997) von Schaewen A., Plant Physiology,
University Osnabrueck, Barbarastr. 11, D-49069 Osnabrueck, GERMANY

FEATURES

Location/Qualifiers

source

1.1975
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/tissue_type="Above ground parts"
71..1801
CDS

BASE COUNT 583 a 382 c 447 g 563 t

ORIGIN

Query Match 22.8%: Score 60; DB 8; Length 1975;

Best Local Similarity 52.5%: Pred. No. 8.1e-06;
Matches 114; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 37 agtgcagaccagaaaggaatggaatgcataatattggaagccattgcttgat 96
DB 728 AGAGCCTCTCAGAGAAAGGCTGGACAGGGTCAATTTGGAAAGCCATTGGCCGTGAC 787
QY 97 gcacttctcccaagctcagacacatatcttccaacttcaggaagaacaata 156
DB 788 TCTGATATCCGGAGACTTACTAGATGCTGAAACAGTACTTACAGAGACCAATC 847
QY 157 tatgaattgncactactagaaaggaatcncaglnaaatccttcacaggtt 216
DB 848 TTCAGATTGATCACTATTGGAAGAGAGCTTGTGGAACCTTCACTGCTCCGATTC 907
QY 217 tcaaanmagnitttgagccacacttngagngnacent 253
DB 908 TCAATCTTGTTCGAGCCCTGTGTGCAAGAAATT 944

RESULT 13

ATG6PDH15

LOCUS 1772 bp mRNA linear PLN 15-AUG-1997
DEFINITION A.thaliana mRNA for glucose-6-phosphate dehydrogenase (clone E15).
ACCESSION X84230

VERSION X84230.1 GI:1174335
KEYWORDS g6pdh gene; glucose-6-phosphate dehydrogenase.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 1772)
Fink,A.F.A.
TITLE Direct Submission
JOURNAL Submitted (23-JAN-1995) A.F.A. Fink, Lab de Biochimie et de
physiologie, Viginales, Universiti de Geneve, 3 Place de
l'Universiti, 1211 Geneva, SWITZERLAND

REFERENCE

AUTHORS

TITLE

Direct Submission
JOURNAL Submitted (23-JAN-1995) A.F.A. Fink, Lab de Biochimie et de
physiologie, Viginales, Universiti de Geneve, 3 Place de
l'Universiti, 1211 Geneva, SWITZERLAND

REFERENCE

AUTHORS

Nucleotide sequence of a cDNA encoding the glucose-6-phosphate
dehydrogenase (GenBank X84230) from Arabidopsis thaliana
Plant Physiol. 108, 1343-1343 (1995)
Ref [3]: Plant Gene Register PGR95-021 (1995).

COMMENT

FEATURES

source

Location/Qualifiers
1.1772
/organism="Arabidopsis thaliana"

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 07:54:40 ; Search time 796.36 Seconds
(without alignments)
567.016 Million cell updates/sec

Title: US-09-300-482-4
Perfect score: 263
Sequence: 1 gaagcatttgatgttcgc.....agnnagacnnnnnnnganna 263

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*

- 1: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61.4	23.3	1485	23	HAAS53321
2	59.6	22.7	6132	21	AACT95684
3	59	22.4	1905	21	AACT95684
4	58.2	22.1	1524	23	AACT95684
5	58.2	22.1	32768	20	AACT95684
6	54	20.5	3015	22	AACT95684
7	54	20.5	3081	22	AACT95684
8	51.4	19.5	1248	23	AACT95684
9	51.4	19.5	1248	23	AACT95684

10	51.4	19.5	1461	23	HAAS51872	Staphylococcus aur
11	51.4	19.5	16592	18	AAV74364	Staphylococcus aur
12	51	19.4	2750	14	AAO38808	G6p coding sequen
13	50.6	19.2	2494	21	AACT95684	Essential Staphylo
14	50.6	19.2	2494	22	AACT95684	Staphylococcus aur
15	50.6	19.2	2494	22	AACT95684	Staphylococcus aur
16	49.6	18.9	1488	19	AACT95684	Staphylococcus aur
17	49.6	18.9	1488	19	AACT95684	Staphylococcus pneu
18	49.6	18.9	1488	19	AACT95684	Staphylococcus pneu
19	48	18.3	2764	19	AACT95684	DNA encoding a S.
20	48	18.3	13121	19	AACT95684	Streptococcus pneu
21	47.4	18.0	2322	24	AACT95684	Mouse ischaemic co
22	44.8	17.0	1488	19	AACT95684	Thermotoga gluco
23	44.8	17.0	10820	20	AACT95684	Polynucleotide seq
24	43	16.3	23532	21	AACT95684	N. meningitidis pa
25	43	16.3	349980	21	AACT95684	Neisseria meningit
26	43	16.3	143768	21	AACT95684	N. meningitidis B
27	42.6	16.2	885	21	AACT95684	Eosinophil activat
28	42.6	16.2	2803	21	AACT95684	Human breast and o
29	42.2	16.0	1482	11	AACT95684	Glucose-6-phosphat
30	42	16.0	3376	22	AACT95684	S. epidermidis gen
31	41.2	15.7	362	21	AACT95684	N. meningitidis pa
32	40.4	15.4	1854	23	AACT95684	Drosophila melanog
33	40.4	15.4	1878	23	AACT95684	Drosophila melanog
34	40.4	15.4	4417	23	AACT95684	Drosophila melanog
35	40.4	15.4	6888	23	AACT95684	Drosophila melanog
36	39.8	15.1	554	20	AACT95684	Polynucleotide seq
37	39.6	15.1	1452	22	AACT95684	C. glutamicum codin
38	39.6	15.1	1665	22	AACT95684	Corynebacterium gl
39	39.6	15.1	2260	18	AACT95684	Brevibacterium fla
40	39.6	15.1	3038	22	AACT95684	C. glutamicum Opca
41	39.6	15.1	6995	22	AACT95684	Nucleotide sequenc
42	39.6	15.1	6995	22	AACT95684	C. glutamicum Opca
43	39.6	15.1	349980	22	AACT95684	C. glutamicum codin
44	39	14.8	780	18	AACT95684	H. pylori cytoplas
45	37.6	14.3	1278	23	AACT95684	Helicobacter pylor

ALIGNMENTS

RESULT 1	HAAS53321	standard; DNA; 1485 BP.
ID	HAAS53321	standard; DNA; 1485 BP.
XX	HAAS53321	
AC	HAAS53321	
XX	HAAS53321	
DF	13-FEB-2002	(first entry)
XX	HAAS53321	
DE	Haemophilus influenzae DNA for cellular proliferation protein #103.	
XX	HAAS53321	
KW	Antisense; ds; prokaryotic cellular proliferation gene;	
KW	antibiotic; antibacterial; drug design.	
XX	HAAS53321	
OS	Haemophilus influenzae.	
XX	HAAS53321	
PN	WO200170955-A2.	
XX	HAAS53321	
PD	27-SEP-2001.	
XX	HAAS53321	
PF	21-MAR-2001; 2001WO-US09180.	
XX	HAAS53321	
PR	21-MAR-2001; 2000US-191078P.	
PR	23-MAY-2000; 2000US-206848P.	
PR	26-MAY-2000; 2000US-207727P.	
PR	23-OCT-2000; 2000US-242578P.	
PR	27-NOV-2000; 2000US-253625P.	
PR	22-DEC-2000; 2000US-257931P.	
XX	16-FEB-2001; 2001US-269308P.	
XX	HAAS53321	
PA	(ELIT-) ELITRA PHARM INC.	
XX	HAAS53321	
PI	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;	

PI Yamamoto RT, Xu HH;
XX
DR WPI: 2001-611495/70.
XX P-PSDB: AAN35462.
XX
PT New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 27: Seq ID No 6958; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1485 BP: 445 A; 274 C; 333 G; 433 T; 0 other;

Query Match 23.3%; Score 61.4; DB 23; Length 1485;
Best Local Similarity 52.1%; Pred. No. 1.8e-08;
Matches 110; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 53 agggatggaatcgcatatatttggaagcatttggttgagacttctccata 112
DB 422 atggctggaagaacgtatctgltgaaaacatttggtatgataaacaagcgcaaa 481
QY 113 ggcctacacaatattcttcaacttcaggaagaacaataatagaattgancctc 172
DB 482 cattagaagctacaatccacgcttctcttggaagacccaattatcgatcgatcact 541
QY 173 tactaggaaggaatcncagttnaaaatccctncaggtttaagggttcaaanmagnittcg 232
DB 542 atttcgtaagaagaacgcttcaaacctgctcgctatgatttttcaaatgttggttg 601
QY 233 agccaccttngagngnagcnnmnganna 263
DB 602 aaccgcttgaacccgtatattcatgatta 632

RESULT 2
AAC79584
ID AAC79584 standard; DNA: 6132 BP.
XX
AC AAC79584;
XX
DT 08-FEB-2001 (first entry)
XX
DE Virulence gene #4.
XX
XX Virulence gene #4.
XX
KW Virulence gene; antibacterial; vaccine; bacterial infection;
KM septicemia; bronchopneumonia; rhinitis; wound infection; ss.
XX
OS Pasteurella multocida.
XX
PN MO200061724-A2.
XX
PD 19-OCT-2000.
XX

PF 06-APR-2000; 2000WO-US09218.
XX
XX 09-APR-1999; 99US-0128689.
PR 10-SEP-1999; 99US-0153453.
XX
XX (PHNA) PHARMACIA & UPJOHN INC.
XX
PI Lowery DE, Fuller TE, Kennedy MJ;
XX
DR WPI: 2000-647422/62.
XX P-PSDB: AAB44524.
XX
PT Attenuated Pasteurellaceae bacteria comprising mutations in virulence
XX genes, useful as a live attenuated vaccine against bacterial infections
XX
PS Claim 1: Pages 72-76; 322pp; English.
XX
CC The family Pasteurellaceae encompasses several pathogens that infect a
CC wide variety of animals. The present invention relates to virulence genes
CC from Pasteurellaceae. The present sequence is one such virulence gene.
CC The present sequence may be mutated in order to produce an inactive gene.
CC The inactive virulence gene may in turn be used to produce a vaccine,
CC which is useful for treating bacterial infections such as septicemias,
CC bronchopneumonias, rhinitis and wound infections.
XX
SQ Sequence 6132 BP: 1838 A; 1154 C; 1409 G; 1731 T; 0 other;

Query Match 22.7%; Score 59.6; DB 21; Length 6132;
Best Local Similarity 51.4%; Pred. No. 9.6e-08;
Matches 110; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 50 agaaggaatgaaatcgcatatatttggaagcatttggttgagacttctcc 109
DB 2903 agtttgctggaagcggttaatttggtgaaaacggttggttatgatacgaagcgcaa 2962
QY 110 atagcctgacaacatattcttcaacttcaggaagaacaataatagaattganc 169
DB 2963 aagaccctcgatattcaaatccacgcttctcttggaagacccaattatcgatcgatgacc 3022
QY 170 atctactaggaaggaatcncagttnaaaatccctncaggtttaagggttcaaanmagnit 229
DB 3023 actactctgtgtaagaagaacggttcaaaaatctgctgtgctgttcttctaagtgatggt 3082
QY 230 ttgagccacttngagngnagcnnmnganna 263
DB 3083 ttgaaccactctggaacgcgaattcatcgatta 3116

RESULT 3
AAC34618
ID AAC34618 standard; DNA: 1905 BP.
XX
AC AAC34618;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 7277.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151920.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156456.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

PR	05-OCT-1999;	990S-0157753.
PR	06-OCT-1999;	990S-0157865.
PR	07-OCT-1999;	990S-0158029.
PR	08-OCT-1999;	990S-0158232.
PR	12-OCT-1999;	990S-0158369.
PR	13-OCT-1999;	990S-0159293.
PR	13-OCT-1999;	990S-0159294.
PR	13-OCT-1999;	990S-0159295.
PR	14-OCT-1999;	990S-0159329.
PR	14-OCT-1999;	990S-0159330.
PR	14-OCT-1999;	990S-0159331.
PR	14-OCT-1999;	990S-0159637.
PR	14-OCT-1999;	990S-0159638.
PR	18-OCT-1999;	990S-0159584.
PR	21-OCT-1999;	990S-0160741.
PR	21-OCT-1999;	990S-0160767.
PR	21-OCT-1999;	990S-0160768.
PR	21-OCT-1999;	990S-0160770.
PR	21-OCT-1999;	990S-0160814.
PR	21-OCT-1999;	990S-0160815.
PR	22-OCT-1999;	990S-0160880.
PR	22-OCT-1999;	990S-0160981.
PR	22-OCT-1999;	990S-0160989.
PR	25-OCT-1999;	990S-0161404.
PR	25-OCT-1999;	990S-0161405.
PR	25-OCT-1999;	990S-0161406.
PR	26-OCT-1999;	990S-0161359.
PR	26-OCT-1999;	990S-0161360.
PR	26-OCT-1999;	990S-0161361.
PR	28-OCT-1999;	990S-0161920.
PR	28-OCT-1999;	990S-0161992.
PR	28-OCT-1999;	990S-0161993.
PR	29-OCT-1999;	990S-0162142.

Query Match	22.4%	Score 59	DB 21	Length 1905
Best Local Similarity	50.4%	Pred. No. 1e-07		
Matches 125	Conservative	0	Mismatches 123	Indels 0
				Gaps 0
QY	6	actttggaatgtgcgtcatgtctcttgcaagcaggtgtctcagaccagaaggaatgaatcg	65	
DB	670	atccgtctgtagtgygttaagtgtgcgaagtcctctagagcctccctaagaaatgygttgacaag	729	
QY	66	cataatatttggagaagccatttggcttctgatactcttcctccctaagctgaacaata	125	
DB	730	ggtcatttggagaagccatttggcttctgatactcttcctccctaagctgaacaata	789	
QY	126	tcttctcttaaaccttccagaaagcaaatatatagaattgancactactagaagaagaa	185	
DB	790	tctgaacaagctatctctacagaggaagcaaatcttcagattgtactaatttggagaaga	849	
QY	186	tencagtnaanaatccnccnaggtttaagggttccaannmagnttttggaccacttngag	245	
DB	850	gcttcttggagaaccttccagtgctccgattccaaactctgttttccgacctctgtgctc	909	
QY	246	mngnaact 253		
DB	910	aagaatt 917		

RESULT 4

AAS52910

ID AAS52910 standard; DNA; 1524 BP.

XX AAS52910;

AC

DT 13-FEB-2002 (first entry)

XX

DE Enterococcus faecalis DNA for cellular proliferation protein #338.

XX

KW Antisense; ds; prokaryotic cellular proliferation gene;

KW antibiotic; antibacterial; drug design.

XX

XX Enterococcus faecalis.

XX WO200170955-A2.
 PN
 XX
 PD 27-SEP-2001.
 XX
 XX 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI: 2001-611495/70.
 DR P-PSDB: AAU35051.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Claim 27; Seq ID No 6547; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antihoddes capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence encodes an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX

Sequence 1524 BP; 484 A; 302 C; 289 G; 449 T; 0 other:

Query Match:	22.1%;	Score	58.2;	DB	23;	Length	1524;
Best Local Similarity	51.2%;	Pred. No.	1.6e-07;				
Matches	108;	Conservative	0;	Mismatches	103;	Indels	0;
				Gaps			0

QY	53	aaggatggaatcgcataatatgttgaagaacattggcttgtagaacattctcccata	112
I			
Db	416	aagcgcttgctgcataatcatcgaagaaccattggcttcgatatagaactgcacag	475
QY	113	ggctgcacaacatactcttcacaacttcacagaaagcaaatatagaattamcatt	172
I			
Db	476	aattaagtgaagaattcgtgcgcattctccgaaccaagatattctccgaattgaccatt	535
QY	173	tactagaagaaggaatcmcagtnaaabctocncaagtttaaggattccaannagnattg	232
I			
Db	536	acttaagcacaagaagaatgatgccaaaatlctcagccattcgtttgcacataatatatttg	595
QY	233	agccacctingagmgnaactnnmnganna	263
I			
Db	596	aatcacacatggaacaaatcgttacattgataa	626

RESULT	5
--------	---

XX	AA13065	
ID	AA13065 standard; DNA: 32768 BP.	
XX		
AC	AA13065;	
XX		
DT	19-MAR-1999 (first entry)	
XX		
DE	Enterococcus faecalis genome contig SEQ ID NO:128.	
KW	Enterococcus faecalis; config; detection: Enterococcal infection; vaccine; attenuation; computer readable medium, ds.	
OS	Enterococcus faecalis.	
XX		
PM	MO9850555-A2.	
PD		
XX	12-NOV-1998.	
PF		
XX	04-MAY-1998; 98WO-US08985.	
PR		
PR	14-NOV-1997; 97US-0066009.	
PR	06-MAY-1997; 97US-0044031.	
XX	16-MAY-1997; 97US-0046655.	
PA	(HUMA-) HUMAN GENOME SCI INC.	
PI	Barash SC, Dillon PJ, Kunsch CA:	
DR		
XX	WPI: 1999-045171/04.	
PT		
PT	New isolated Enterococcus faecalis polynucleotides and polypeptides - used to develop products for the detection of Enterococcus and for use in vaccines for prevention or attenuation of Enterococcus infection.	
PS		
XX	Claim 1: Page 767-783; 2084pp: English.	
CC	A computer readable medium has been developed which has recorded on it 982 nucleotide sequences isolated from the Enterococcus faecalis genome. CC AA12938 to AA13919 represent these nucleotide sequences which are CC primary nucleotide sequences, also known as contigs. The computer-based CC system can identify fragments of the Enterococcus faecalis genome with CC commercial importance. The products can be used to detect the presence CC of Enterococcus faecalis in samples. They can also be used for CC diagnosing Enterococcal infection in an animal and monitoring CC progression of disease, and for identifying agents which can be used to CC modulate the growth or pathogenicity of Enterococcus faecalis, or CC another related organism, in vivo or in vitro. In particular the CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences CC can be used in vaccines to prevent or attenuate an Enterococcal CC infection.	
SQ	Sequence 32768 BP; 9529 A; 6883 C; 5611 G; 10720 T; 25 other:	
Query Match	22.1%; Score 58.2; DB 20; Length 32768;	
Best Local Similarity	51.2%; Pred. No. 4.2e-07;	
Matches 108; Conservative	0; Mismatches 103; Indels 0; Gaps 0,	
OY	53 aggaatggaatcgcaataatttggaaagccattgttgctttgaagcaactttcaccata 112	
DB	19293 aaggttcgatgcgtctaattatcatcagaagccatttgttcgatgatgatcgtccta 19352	
OY	113 ggcgtcacacaatatcttcttccaaccttcaggaaaagaacaaatatagaaatgancatc 172	
DB	19353 aatttaagaagaataatcgtgtcgtatcttcctgaacaagatatlttccgaattgacatt 19412	
OY	173 tactaggaaggaatencagtnaaaatccnccaggtttaagggttccaannmagntttg 232	
DB	19413 acttaggcacaagaataatgatccaaatatattccagccattcgttttgcacaataattttg 19472	
OY	233 agccaccttgaggnnacnlmnnngana 263	

Query Match	Best Local Similarity	Score 54:	DB 22:	Length 3015:
Matches 114; Conservative	54.5%;	Pred. NO. 3.5e-06;		
	0;	Mismatches 93;	Indels 2;	Gaps 1;
Db	1315	gaccgaagaaggatggaatcgcataataattagaagccattggcttgatgcatttc	104	
QY	45	gaccgaagaaggatggaatcgcataataattagaagccattggcttgatgcatttc	104	
Db	1315	gaacgtaaaggaggatgcatggaataattatagaagaacacatttggtgatgttaataac	1374	
QY	105	ttccatagatgctgacacaaatatcttccttcaaacatttcaggaagaagcaatatagaat	164	
Db	1375	t--gcataaataatataataataagaagaagatctctttaaagaagaataatatttggaat	1432	

PI Yamamoto RT, Xu HH;
 XX
 DR WPI: 2001-611495/70.
 XX P-PSDB; AAU34013.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX
 PS Claim 27: Seq ID No 4454; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence encodes an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1461 BP; 521 A; 231 C; 257 G; 452 T; 0 other:
 XX
 Query Match 19.5%; Score 51.4; DB 23; Length 1461;
 Best Local Similarity 51.0%; Pred. No. 1.6e-05;
 Matches 103; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
 XX
 QY 52 aaggatggaatcgatcataattgagagcattgcttgatgcacttctccat 111
 DB 415 aaagatttaaacgctctgtatcgaataaccatcgtagtatttaaacagcgaa 474
 QY 112 aggcgtacacatatctcttcaacttcaggaagaacaatatagattgancat 171
 DB 475 gcaatcaacatcaatcgttaaatcatttaagaagaagaattatcgtatcgaccac 534
 QY 172 ctactagaagaagatcncagtnaaatcctcncagggttaagggttccaannnagnttt 231
 DB 535 tatttaagaaaagacatggttcaaatatcgcaggtatcttgcgaatgcgagtgtt 594
 QY 232 gagcaccttngagngnact 253
 DB 595 gaaccattatgataacaat 616
 XX
 RESULT 11
 AAV74364
 ID AAV74364 standard; DNA; 16592 BP.
 AC AAV74364;
 XX
 DT 16-MAR-1999 (first entry)
 XX
 DE *Staphylococcus aureus* contig SEQ ID #53.
 XX
 KW Computer readable medium: vaccine; *S.aureus* infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.
 XX
 OS *Staphylococcus aureus*.
 XX
 FH key Location/Qualifiers
 FT misc_feature 121..180

FT /*tag- a
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT 1921..1980
 FT /*tag- b
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT 3721..3780
 FT /*tag- c
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT 5521..5580
 FT /*tag- d
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT 7321..7380
 FT /*tag- e
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT 9121..9180
 FT /*tag- f
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT 10921..10980
 FT /*tag- g
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT 12721..12780
 FT /*tag- h
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT 14521..14580
 FT /*tag- i
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT 16321..16380
 FT /*tag- j
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT misc_feature
 FT 16321..16380
 FT /*tag- j
 FT /note- "these bases represent a line of missing text in
 FT the sequence listing in the specification. They
 FT are included to maintain the nucleotide numbering
 FT given in the specification for this DNA sequence"
 FT EP786519-A2.
 PD 30-JUL-1997.
 XX
 XX 07-JAN-1997: 97EP-0100117.
 PF
 XX 05-JAN-1996: 96US-0009861.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
 PI Rosen CA;
 XX

DR WPI: 2000-282222/24.
XX Novel methods for screening for antibacterial agents, useful for
PT treatment or prevention of Staphylococcus infection, by testing
PT compounds against essential bacterial genes -
XX
PS Disclosure: Column 35-38; 256pp; English.
XX
CC The present sequence represents an essential Staphylococcus aureus gene.
CC The invention relates to a method for screening for an antibacterial
CC agent. The method comprises determining if a test compound is active
CC against the essential bacterial genes (AA26850-A26956). The invention
CC also includes a method for evaluating an agent for activity on the
CC essential genes. Also included in the invention is the production of an
CC antibacterial agent. The antibacterial agents of the invention are used
CC to treat or prevent bacterial infections, particularly where caused by
CC Staphylococcus aureus. The antibacterial agents are unlikely to be
CC affected by known resistance mechanisms.
XX
SQ Sequence 2494 BP; 809 A; 396 C; 389 G; 880 T; 20 other;

Query Match 19.2%; Score 50.6; DB 21; Length 2494;
Best Local Similarity 50.0%; Pred. No. 3.4e-05;
Matches 101; Conservative 2; Mismatches 99; Indels 0; Gaps 0;
QY 52 aaggatggaatcgataataattgagaagccattggtgcttgatgcattcttcccat 111
|| |||| || |||| || || || || || || || || || || || || || || || ||
DB 1425 AAAGATTTAAACGCGCTGTATCGAAAACATTGCGTAGTGATTTAAATCAGCCGAA 1366
QY 112 aggcgcacacatatctcttcacaaacttcagaaagcaaatatagaaatgancat 171
|| |||| || |||| || || || || || || || || || || || || || || || ||
DB 1365 GCATTAAACAATCAAAATTCGTAATCATTTAAAGAAAGAAATTTAVCGTATKGACCAC 1306
QY 172 ctactaggaagaatcncagtnaaatccctncaggtttaaggtttcaaanngntttt 231
|| |||| || |||| || || || || || || || || || || || || || || || ||
DB 1305 TATTAGGAAAGACATGCTTCAAAATATCGAGTATTACGTTTCCGAAATCGATGTTT 1246
QY 232 gaggcaccttngagngnancat 253
|| |||| || || || || || || || || || || || || || || || || || || ||
DB 1245 GAACCATTTATGGAATTAACAAT 1224

RESULT 14
AAS08003/C
ID AAS08003 standard; DNA: 2494 BP.
XX
AC AAS08003;
XX
DT 23-OCT-2001 (first entry)
XX
DE Staphylococcus aureus essential pathogenic bacterial DNA #3.
XX
KW Essential bacterial gene; antibacterial agent; drug screening; ds.
XX
OS Staphylococcus aureus.
XX
PN US6228588-B1.
XX
PD 08-MAY-2001.
XX
PF 09-MAR-1999; 99US-0266417.
XX
PR 15-SEP-1995; 95US-0003798.
XX
PR 22-DEC-1995; 95US-0009102.
XX
PR 13-SEP-1996; 96US-0714918.
XX
PA (MICR-) MICROCID PHARM INC.
XX
PI Benton B, Lee VJ, Malouin F, Martin PK, Schmid MB, Sun D;
XX
DR WPI: 2001-342669/36.
XX

PT Screening or producing antibacterial agents active against
PT Staphylococcus aureus target genes, comprises contacting a cell or a
PT test compound with a gene product or polypeptide that is encoded by
PT essential bacterial genes -
XX
PS Claim 7: Column 35-37; 251pp; English.
XX
CC The sequence represents a DNA encoding a Staphylococcus aureus essential
CC pathogenic bacterial protein. These polynucleotides are useful for
CC screening for and producing antibacterial agents against pathogenic
CC bacteria. A polypeptide that is encoded by an essential bacterial gene
CC can be contacted with a cell or a test compound and the level of activity
CC against the bacterial protein is monitored. The method is used in drug
CC discovery and drug production, particularly for screening and preparing
CC compounds active against Staphylococcus aureus target genes. These
CC compounds are useful for inhibiting the growth of pathogenic bacteria.
XX
SQ Sequence 2494 BP; 809 A; 396 C; 389 G; 880 T; 20 other;

Query Match 19.2%; Score 50.6; DB 22; Length 2494;
Best Local Similarity 50.0%; Pred. No. 3.4e-05;
Matches 101; Conservative 2; Mismatches 99; Indels 0; Gaps 0;
QY 52 aaggatggaatcgataataattgagaagccattggtgcttgatgcattcttcccat 111
|| |||| || |||| || || || || || || || || || || || || || || || ||
DB 1425 AAAGATTTAAACGCGCTGTATCGAAAACATTGCGTAGTGATTTAAATCAGCCGAA 1366
QY 112 aggcgcacacatatctcttcacaaacttcagaaagcaaatatagaaatgancat 171
|| |||| || |||| || || || || || || || || || || || || || || || ||
DB 1365 GCATTAAACAATCAAAATTCGTAATCATTTAAAGAAAGAAATTTAVCGTATKGACCAC 1306
QY 172 ctactaggaagaatcncagtnaaatccctncaggtttaaggtttcaaanngntttt 231
|| |||| || |||| || || || || || || || || || || || || || || || ||
DB 1305 TATTAGGAAAGACATGCTTCAAAATATCGAGTATTACGTTTCCGAAATCGATGTTT 1246
QY 232 gaggcaccttngagngnancat 253
|| |||| || || || || || || || || || || || || || || || || || || ||
DB 1245 GAACCATTTATGGAATTAACAAT 1224

RESULT 15
AAF91534/C
ID AAF91534 standard; DNA: 2494 BP.
XX
AC AAF91534;
XX
DT 10-MAY-2001 (first entry)
XX
DE Staphylococcus aureus essential gene #3.
XX
KW Screening; antibacterial; infection; ds.
XX
OS Staphylococcus aureus.
XX
PN US6187541-B1.
XX
PD 13-FEB-2001.
XX
PF 09-MAR-1999; 99US-0265315.
XX
PR 15-SEP-1995; 95US-0003798.
XX
PR 22-DEC-1995; 95US-0009102.
XX
PR 13-SEP-1996; 96US-0714918.
XX
PA (MICR-) MICROCID PHARM INC.
XX
PI Benton B, Lee VJ, Malouin F, Martin PK, Schmid MB, Sun D;
XX
DR WPI: 2001-158814/16.
XX
PT Novel methods for screening for antibacterial agents, useful for
PT treatment or prevention of Staphylococcus infection, by testing

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 07:40:58 ; Search time 185.98 Seconds
(without alignments)
347.358 Million cell updates/sec

Title: US-09-300-482-4

Perfect score: 263

Sequence: 1 gaagacatttgatgttgc.....agunnaatnnnnnganna 263

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCUS.COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	50.6	19.2	2494	3	US-08-714-918-3 Sequence 3, Appli
C 2	50.6	19.2	2494	4	US-09-265-315-3 Sequence 3, Appli
C 3	50.6	19.2	2494	4	US-09-265-315-3 Sequence 3, Appli
C 4	50.6	19.2	2494	4	US-09-265-315-3 Sequence 3, Appli
C 5	49.6	18.9	1488	3	US-08-962-859A-1 Sequence 1, Appli
C 6	48	18.3	2764	2	US-08-962-859A-1 Sequence 1, Appli
7	37.4	14.2	1461	3	US-08-445-463B-5 Sequence 5, Appli
8	37.4	14.2	1461	3	US-08-445-463B-5 Sequence 5, Appli
9	37.4	14.2	1461	5	PCT-US94-03437-5 Sequence 5, Appli
10	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
11	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
12	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
13	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
14	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
15	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
16	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
17	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
18	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
19	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
20	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
21	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
22	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
23	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
24	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
25	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
26	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli
27	37.2	14.1	1455	5	PCT-US94-03437-7 Sequence 7, Appli

28	29.8	11.3	1506	3	US-08-445-463B-1 Sequence 1, Appli
29	29.8	11.3	1506	3	US-08-445-463B-1 Sequence 1, Appli
30	29.8	11.3	1506	3	US-08-445-463B-1 Sequence 1, Appli
31	29.8	11.3	1506	3	US-08-445-463B-1 Sequence 1, Appli
32	29.8	11.3	1506	3	US-08-445-463B-1 Sequence 1, Appli
33	29.6	11.3	1484	4	US-09-210-168-2 Sequence 2, Appli
34	29.6	11.3	1517	1	US-08-032-869A-1 Sequence 1, Appli
35	29.6	11.3	1517	1	US-08-032-869A-1 Sequence 1, Appli
36	29.6	11.3	1517	1	US-08-472-673-1 Sequence 1, Appli
37	29.6	11.3	1517	1	US-08-472-673-1 Sequence 1, Appli
38	29.6	11.3	1517	1	US-08-472-673-1 Sequence 1, Appli
39	29.6	11.3	1517	2	US-08-472-678-2 Sequence 2, Appli
40	29.6	11.3	1517	2	US-08-472-678-2 Sequence 2, Appli
41	29.6	11.3	1517	2	US-08-472-678-2 Sequence 2, Appli
42	29.6	11.3	1517	2	US-08-472-678-2 Sequence 2, Appli
43	29.6	11.3	1518	4	US-08-944-368A-1 Sequence 1, Appli
44	29.6	11.3	1518	4	US-08-944-368A-1 Sequence 1, Appli
45	29.6	11.3	7904	1	US-08-316-229B-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-714-918-3/c
Sequence 3, Application US/08714918
Patent No. 6037123
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714.918
FILING DATE: September 13, 1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 232/005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2494 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

ADDRESS: Lyon & Lyon
 STREET: 633 West Fifth Street
 STREET: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071-2066
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/265,315
 FILING DATE: March 9, 1999
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/714,918
 FILING DATE: September 13, 1996
 APPLICATION NUMBER: 60/009,102
 FILING DATE: December 22, 1995
 APPLICATION NUMBER: 60/003,798
 FILING DATE: September 15, 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Warburg, Richard J.
 REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 240/247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2494 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-265-315-3

Query Match 19.2%; Score 50.6; DB 4; Length 2494;
Best Local Similarity 50.0%; Pred. No. 7.2e-07;
Matches 101; Conservative 2; Mismatches 99; Indels 0; Gaps 0;

QY 52 aaggatggaatcgacataatttgagaagccattggcttgatgcacttctcccat 111
DB 1425 AAAGATTAAACGCTTGTATCGAAAAACCATTCGGTAGATTAAATCAGCCGAA 1366
QY 112 agctgcacacatactcttccaacttcaggaaaagcaatatagattgancat 171
DB 1365 GCATTAAACATCAAAATTCGTAATCATTTAAAGAGACAAATTTAYCGTATKGCAC 1306
QY 172 ctactagaagaatcncaglnaaatcctncaggtttaagggtttcaaanmagnltt 231
DB 1305 TATTAGAAAGACATGTTCAAAATATCGAGGTATTACGTTTGGCAATGCGATGTTT 1246
QY 232 gaggcaacttngagmgnaent 253
DB 1245 GAACCATTTAGTAATCAAAAT 1224

RESULT 4
US-09-266-417-3/c
Sequence 3, Application US/09266417
Patent No. 6228588
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving J.
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/266,417
FILING DATE: March 9, 1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/714,918
FILING DATE: September 13, 1996
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995

APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Marburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 240/248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2494 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-266-417-3

Query Match 19.2%; Score 50.6; DB 4; Length 2494;
Best Local Similarity 50.0%; Pred. No. 7.2e-07;
Matches 101; Conservative 2; Mismatches 99; Indels 0; Gaps 0;

QY 52 aaggatggaatcgacataatttgagaagccattggcttgatgcacttctcccat 111
DB 1425 AAAGATTAAACGCTTGTATCGAAAAACCATTCGGTAGATTAAATCAGCCGAA 1366
QY 112 agctgcacacatactcttccaacttcaggaaaagcaatatagattgancat 171
DB 1365 GCATTAAACATCAAAATTCGTAATCATTTAAAGAGACAAATTTAYCGTATKGCAC 1306
QY 172 ctactagaagaatcncaglnaaatcctncaggtttaagggtttcaaanmagnltt 231
DB 1305 TATTAGAAAGACATGTTCAAAATATCGAGGTATTACGTTTGGCAATGCGATGTTT 1246
QY 232 gaggcaacttngagmgnaent 253
DB 1245 GAACCATTTAGTAATCAAAAT 1224

RESULT 5
US-08-962-859A-1
Sequence 1, Application US/08962859A
Patent No. 6127345
GENERAL INFORMATION:
APPLICANT: Burnham, Martin K. R.
TITLE OF INVENTION: No. 6127345el Glucose 6-Phosphate
TITLE OF INVENTION: Dehydrogenase Gene
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/962,859A
FILING DATE: 03-NOV-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/035,072
FILING DATE: 21-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Jackson, Arthur E
REGISTRATION NUMBER: 34,354
REFERENCE/DOCKET NUMBER: GMS0001
TELECOMMUNICATION INFORMATION:

Search completed: July 3, 2002, 07:41:01
Job time: 24026 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 07:54:49 ; Search time 796.36 Seconds
(without alignments)
883.941 Million cell updates/sec

Title: US-09-300-482-14

Perfect score: 410

Sequence: 1 cccacgcgtccgcgcgtatg.....gcaacgcgcgccttcctgatg 410

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 segs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.*
23: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2003.DAT.*
24: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2004.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	260.4	63.5	1753	21	Arabidopsis thaliana
2	243	59.3	483	21	Zea mays DNA fragm
3	242.4	59.1	466	21	Zea mays DNA fragm
4	195	47.6	477	21	Zea mays DNA fragm
5	175	42.7	1639	21	Arabidopsis thaliana
6	173.4	42.3	1639	21	Arabidopsis thaliana
7	168	41.0	369	21	Zea mays DNA fragm
8	99.2	24.2	1916	18	Brevibacterium fli
9	97.6	23.8	1476	22	C glutamicum codin

10	97.6	23.8	1537	22	AAE71486
11	97.6	23.8	1575	22	AAE71485
12	97.6	23.8	2335	22	AAI65505
13	97.6	23.8	349980	22	AAH68528
14	97.6	23.8	349980	22	AAH68529
15	94.4	23.0	1508	23	ABL06613
16	91.2	22.2	1038602	20	AAZ01425
17	89.8	21.9	1455	23	AAE53320
18	88	21.5	273254	21	AAH81914
19	85.8	20.9	1458	22	AAH52008
20	83.8	20.4	2310	24	AAE94752
21	80.2	19.6	92834	21	AAH81473
22	80.2	19.6	349980	21	AAE21544
23	69.2	16.9	14516	20	AAH06748
24	69.2	16.9	14516	21	AAZ56375
25	68.8	16.8	1419	23	AAE51371
26	68.8	16.8	1422	23	AAE52744
27	67.2	16.4	29912	23	AAE55607
28	66.2	16.1	4487	7	AAH60943
29	65.6	16.0	2450	20	AAH20537
30	62.8	15.3	1407	22	AAH53136
31	62.8	15.3	3516	22	AAH54169
32	62.8	15.3	3773	22	AAH54437
33	61.2	14.9	1407	23	AAE56087
34	58	14.1	1407	21	AAE52924
35	58	14.1	1407	21	AAE52925
36	58	14.1	1407	21	AAE52926
37	58	14.1	1407	21	AAE52927
38	58	14.1	1407	21	AAE52929
39	58	14.1	1407	21	AAE52930
40	58	14.1	1407	21	AAE52935
41	58	14.1	1407	21	AAE52936
42	57	13.9	2025	19	AAV42965
43	57	13.9	2025	20	AAH16683
44	57	13.9	2025	20	AAH16683
45	56.4	13.8	1407	21	AAE52931

ALIGNMENTS

RESULT 1	
AAC37505	standard; DNA; 1753 BP.
ID AAC37505;	
AC AAC37505;	
XX	
XX	
DT 17-OCT-2000	(first entry)
XX	
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 17627.	
XX	
KW Hybridisation assay; genetic mapping; gene expression control;	
KW Protein identification; signal transduction pathway;	
KW metabolic pathway; promoter; termination sequence; ss.	
XX	
OS Arabidopsis thaliana.	
XX	
PN EP1033405-A2.	
XX	
PD 06-SEP-2000.	
XX	
PF 25-FEB-2000; 2000EP-0301439.	
XX	
PR 05-FEB-1999; 99US-01231825.	
PR 05-MAR-1999; 99US-0123180.	
PR 09-MAR-1999; 99US-0123548.	
PR 23-MAR-1999; 99US-0125788.	
PR 25-MAR-1999; 99US-0126264.	
PR 29-MAR-1999; 99US-0126785.	
PR 01-APR-1999; 99US-0127462.	
PR 06-APR-1999; 99US-0128234.	
PR 08-APR-1999; 99US-0128714.	
PR 16-APR-1999; 99US-0129845.	

Corynebacterium gl
Corynebacterium gl
Nucleotide sequenc
C glutamicum codin
C glutamicum codin
Drosophila melanog
Complete genome se
Haemophilus influ
Mycobacterium tube
Human DNA sequenc
N. meningitidis pa
Neisseria meningit
E. coli O111 antiG
Escherichia coli f
Enterococcus faeca
Enterococcus faeca
Protonibacterium
Sequence encoding
Polynucleotide seq
S. epidermidis ope
S. epidermidis gen
S. epidermidis gen
Salmonella typhi D
Escherichia coli 0
AAE52924
AAE52925
Escherichia coli 0
Escherichia coli 0
Escherichia coli 0
Escherichia coli 0
Escherichia coli 0
Escherichia coli 0
E. coli DNA for ce
Streptococcus pneu
Streptococcus pneu
Escherichia coli 0

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151308.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154038.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 63.5%; Score 260.4; DB 21; Length 1753;
Best Local Similarity 78.2%; Pred. No. 2,6e-58;
Matches 312; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 12 ggcgtatcatggcagagaacttgcctcaacatgacagagaaggtcccatctgtg 71
DB 146 gctgtgataggcagagaacttgcctcaacatgacagagaaggtcccatctgtt 205
QY 72 tacaacaggaacaactccaaggtgagacagacgltgcaggtgccaaaggcagaaggaac 131
DB 206 tacaacaggaactacttccaaggttgatgacacgltgcagacagcacaagaaggcaat 265
QY 132 ctcccgcttaagcgtctcaatgaaccccgctccttgttaagttcaattcagaagccag 191
DB 266 ctccctcttaagcgtctcaatgaaccccgagctcttgttaagttcaattcagaagccag 325
QY 192 gtggtatcatgctgcagacccgcgcgcagctgacacagaacatgcgcagccgca 251
DB 326 gttatatacatcgttgtaagcgtgttctccggtgacacagccatcagaccctctct 385
QY 252 gctcaacttgagcagagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 311
DB 386 gcttatttgagagaagagtgatgttattgtatgtgtgtaataagatgcatgagaact 445
QY 312 gagagagagagagaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 371
DB 446 gagagagagagaagacccgtgagcagagacggttccctcacttgtaattgaggttct 505
QY 372 gagagagaagagaggtggtccgcagacggtccgtcgtatg 410
DB 506 ggtgtgtgagaaggtgtcgtcgtatggtccatcatgagt 544

RESULT 2

AAC38393
ID AAC38393 standard; DNA; 483 BP.

XX AAC38393;

DT 17-OCT-2000 (first entry)

DE zea mays DNA fragment SEQ ID NO: 20807.

XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
XX pathway; promoter; termination sequence; corn; ss.
XX zea mays subsp. mays.
OS

XX XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136788.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.

ID AAC41032 standard; DNA; 466 BP.
XX AAC41032;
XX
XX
DT 17-OCT-2000 (first entry)
XX
DE Zea mays DNA fragment SEQ ID NO: 30405.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 28-MAY-1999; 990S-0136392.
PR 01-JUN-1999; 990S-0136782.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0138492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0138459.
PR 18-JUN-1999; 990S-0139460.

PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 21-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145818.
PR 27-JUL-1999; 990S-0145819.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0148723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162442.

Query Match 59.1%; Score 242.4; DB 21; Length 466;
Best Local Similarity 99.2%; Pred. No. 9e-54;
Matches 254; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 12 ggggtatgtgggagaaacctgtccctcaaatgtcagagaagaaggttcccatctctgtg 71
DB 212 ggggtatgtgggagaaacctgtccctcaaatgtcagagaagaaggttcccatctctgtg 271
OY 72 tacaacagacaacacctcaaggttgagcagacgtgcagcgtgccaaaggcagaaggaac 131
DB 272 tacaacagacaacacctcaaggttgagcagacgtgcagcgtgccaaaggcagaaggaac 331
OY 132 ctcccgcttaagcgttccatgaaccgcgtctcttgaagtccaattcagaagccagcg 191
DB 332 ctcccgcttaagcgttccatgaaccgcgtctcttgaagtccaattcagaagccagcg 391
OY 192 gtgggagatattctcgttcaagcgcgcgccaggttaccacaacatcgcgagctcgca 251

DB 392 gtgggtatcatgtctgtcgaagccgcgcgcacagttgaccaga-catcgcgacgtcgca 450
OY 252 gctcaacttgagcagcg 267
DB 451 gctcaacttgagcagcg 466
RESULT 4
AAC35766 standard; DNA: 477 BP.
ID AAC35766 standard; DNA: 477 BP.
AC AAC35766;
XX
XX
DT 17-OCT-2000 (first entry)
DE
DE Zea mays DNA fragment SEQ ID NO: 11344.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.


```
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159284.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 47.6%; Score 195; DB 21; Length 477;
Best Local Similarity 86.9%; Pred. No. 2e-41;
Matches 213; Conservative 1; Mismatches 31; Indels 0; Gaps 0;

QY 11 cgcgctatgggagagaaccttgcctcaaatctgcagagaagggttcccatctgt 70
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||
Db 233 cgcgctatggagagaaccttgcctcaaatctgcgagagaagggttcccatctcgt 292
QY 71 gtacaacaggaacaacctccaagtgtgacgagacccgtgcagcgtgcacaaaggcagaaga 130
```


Db 293 ctacacacgagacacctcaaggtgatgagacgctgacgctgcacaggtcgaagga 352
QY 131 cctccgcctcaagcctccatgaccgcgctcctcttgtagatccatccagaagca 190
Db 353 cctccgcctgcttggtctccacgaccgcgctcctcttgtagctccacgaagcccc 412
QY 191 ggtggtgcatcgtctcgtcaagcgcgccgacgttaccagacatcgacgctcgc 250
Db 413 tgcgtcatcatgctcgtcgaagcgtggcgccgctgtagcacagacatgtcgaagctcgc 472
QY 251 agctc 255
Db 473 ggcgc 477

RESULT 5
AAC46021
ID AAC46021 standard; DNA: 1639 BP.
XX AAC46021;
AC AAC46021;
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 48624.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123348.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128334.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136592.
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

PR	05-AUG-1999	9905-0147199
PR	05-AUG-1999	9905-0147260
PR	06-AUG-1999	9905-0147303
PR	06-AUG-1999	9905-0147416
PR	06-AUG-1999	9905-0147463
PR	09-AUG-1999	9905-0147535
PR	09-AUG-1999	9905-0147871
PR	11-AUG-1999	9905-0148319
PR	12-AUG-1999	9905-0148341
PR	13-AUG-1999	9905-0148365
PR	13-AUG-1999	9905-0148368
PR	16-AUG-1999	9905-0149384
PR	17-AUG-1999	9905-0149175
PR	18-AUG-1999	9905-0149426
PR	20-AUG-1999	9905-0149122
PR	20-AUG-1999	9905-0149723
PR	20-AUG-1999	9905-0149929
PR	22-AUG-1999	9905-0149902
PR	23-AUG-1999	9905-0149930
PR	25-AUG-1999	9905-0150066
PR	26-AUG-1999	9905-0150084
PR	27-AUG-1999	9905-0150185
PR	27-AUG-1999	9905-0151060
PR	27-AUG-1999	9905-0151086
PR	30-AUG-1999	9905-0151303
PR	31-AUG-1999	9905-0151438
PR	01-SEP-1999	9905-0151930
PR	07-SEP-1999	9905-0152363
PR	10-SEP-1999	9905-0153070
PR	13-SEP-1999	9905-0153758
PR	15-SEP-1999	9905-0154039
PR	16-SEP-1999	9905-0154739
PR	20-SEP-1999	9905-0154779
PR	22-SEP-1999	9905-0155139
PR	24-SEP-1999	9905-0155486
PR	24-SEP-1999	9905-0155569
PR	28-SEP-1999	9905-0156458
PR	29-SEP-1999	9905-0156596
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157753
PR	06-OCT-1999	9905-0157865
PR	07-OCT-1999	9905-0158032
PR	08-OCT-1999	9905-0158232
PR	12-OCT-1999	9905-0158369
PR	13-OCT-1999	9905-0158935
PR	13-OCT-1999	9905-0159294
PR	13-OCT-1999	9905-0159295
PR	14-OCT-1999	9905-0159529
PR	14-OCT-1999	9905-0159530
PR	14-OCT-1999	9905-0159931
PR	14-OCT-1999	9905-0159637
PR	14-OCT-1999	9905-0159638
PR	18-OCT-1999	9905-0159584
PR	21-OCT-1999	9905-0160741
PR	21-OCT-1999	9905-0160747
PR	21-OCT-1999	9905-0160768
PR	21-OCT-1999	9905-0160770
PR	21-OCT-1999	9905-0160814
PR	21-OCT-1999	9905-0160815
PR	22-OCT-1999	9905-0160980
PR	22-OCT-1999	9905-0160981
PR	22-OCT-1999	9905-0160989
PR	25-OCT-1999	9905-0161404
PR	25-OCT-1999	9905-0161405
PR	25-OCT-1999	9905-0161406
PR	26-OCT-1999	9905-0161359
PR	26-OCT-1999	9905-0161360
PR	26-OCT-1999	9905-0161361
PR	28-OCT-1999	9905-0161920
PR	28-OCT-1999	9905-0161922
PR	28-OCT-1999	9905-0161993
PR	28-OCT-1999	9905-0161994
PR	29-OCT-1999	9905-0162143

Query Match	42.7%	Score 175	DB 21	Length 1639	
Best Local Similarity	64.8%	Pred. No. 4,4e-36			
Matches 259	Conservative 0	Mismatches 141	Indels 0	Gaps 0	
QY	11	cgcgcgtcatggcgagcaaccttcgcccctaacaattgcagagaagggctcccatctcgt	70		
DB	101	cgcagtcattggacaacaaccttcgctttaacctgcgcgataaagattcccaatctcgt	160		
QY	71	gtacacagagaaacacctcaaggttggacgagaccgctgcagcgtgcacaaagcgagaagaa	130		
DB	161	ctacaatcgaacctctcccaagaagtcgacgaacaccttagatcgtcgtcccaacgaaggaa	220		
QY	131	ccttcgcgtctcggcgtctccatgatgaccccggtccctttgtgaagtccattcgaagccag	190		
DB	221	actcccaatcgctgtgtgtcctaatactcgtcgtcgatttctctctcgtatccaacggtctag	280		
QY	191	ggtgtgtatcatcgtcgtctcaagcgcgcgcgagttgacagacatcgcagcgtcgc	250		
DB	281	atccgttatcatcctcttgcataaagccggtgctccgcttgacccaacatcctcgtccttc	340		
QY	251	agcttacttggacagggcgacgtcatgatcgcattgaggggaaacgagttggtacgagaacac	310		
DB	341	tgaatacatcgtgagcctgtgtgattgtatcatcgcagcgttgaaatgagttgatcagaacac	400		
QY	311	ggagagagagggaggaagcccatgtgaggagcgcgccctnctgtatcttgcatgtgtctc	370		
DB	401	agagcgacgaatcgtcttgaagctcgacaagaagaagattgcttatttagtatgtggaagctc	460		
QY	371	tggaggaagaggggtgcccgcacaagccgcgtccttgatg	410		
DB	461	cgggtgtgagaagaagagctcgtaatggtctctctctatg	500		
RESULT	6				
AAC39123					
ID	AAC39123	standard	DNA	1639	BP.
XX					
AC	AAC39123				
XX					
DT	17-OCT-2000	(first entry)			
XX					
DE	Arabidopsis thaliana	DNA fragment SEQ ID NO: 23456.			
XX					
KM	Hybridisation assay; genetic mapping; gene expression control;				
KW	protein identification; signal transduction pathway;				
KW	metabolic pathway; promoter; termination sequence; ss.				
XX					
OS	Arabidopsis thaliana.				
XX					
XX	EP1033405-A2.				
PN					
PD	06-SEP-2000.				
XX					
PF	25-FEB-2000	2000EP-0301439.			
XX					
PR	25-FEB-1999	99US-0121825.			
PR	05-MAR-1999	99US-0123180.			
PR	09-MAR-1999	99US-0123548.			
PR	23-MAR-1999	99US-0125788.			
PR	25-MAR-1999	99US-0126264.			
PR	29-MAR-1999	99US-0126785.			
PR	01-APR-1999	99US-0127462.			
PR	06-APR-1999	99US-0128234.			
PR	08-APR-1999	99US-0128714.			
PR	16-APR-1999	99US-0129845.			
PR	19-APR-1999	99US-0130077.			
PR	21-APR-1999	99US-0130449.			
PR	23-APR-1999	99US-0130510.			
PR	23-APR-1999	99US-0130891.			
PR	28-APR-1999	99US-0131448.			
PR	30-APR-1999	99US-0132048.			
PR	30-APR-1999	99US-0132407.			
PR	04-MAY-1999	99US-0132484.			

PR 05-MAY-1999;	99US-0132485.
PR 06-MAY-1999;	99US-0132486.
PR 06-MAY-1999;	99US-0132487.
PR 07-MAY-1999;	99US-0132863.
PR 11-MAY-1999;	99US-0134226.
PR 14-MAY-1999;	99US-0134218.
PR 14-MAY-1999;	99US-0134219.
PR 14-MAY-1999;	99US-0134221.
PR 14-MAY-1999;	99US-0134370.
PR 18-MAY-1999;	99US-0134768.
PR 19-MAY-1999;	99US-0134941.
PR 20-MAY-1999;	99US-0135124.
PR 21-MAY-1999;	99US-0135353.
PR 24-MAY-1999;	99US-0135629.
PR 25-MAY-1999;	99US-0136021.
PR 27-MAY-1999;	99US-0136392.
PR 28-MAY-1999;	99US-0136782.
PR 01-JUN-1999;	99US-0137222.
PR 03-JUN-1999;	99US-0137528.
PR 04-JUN-1999;	99US-0137502.
PR 07-JUN-1999;	99US-0137724.
PR 08-JUN-1999;	99US-0138094.
PR 10-JUN-1999;	99US-0138540.
PR 10-JUN-1999;	99US-0138847.
PR 14-JUN-1999;	99US-0139119.
PR 16-JUN-1999;	99US-0139452.
PR 16-JUN-1999;	99US-0139453.
PR 17-JUN-1999;	99US-0139492.
PR 18-JUN-1999;	99US-0139454.
PR 18-JUN-1999;	99US-0139455.
PR 18-JUN-1999;	99US-0139456.
PR 18-JUN-1999;	99US-0139457.
PR 18-JUN-1999;	99US-0139458.
PR 18-JUN-1999;	99US-0139459.
PR 18-JUN-1999;	99US-0139460.
PR 18-JUN-1999;	99US-0139461.
PR 18-JUN-1999;	99US-0139462.
PR 18-JUN-1999;	99US-0139463.
PR 18-JUN-1999;	99US-0139750.
PR 21-JUN-1999;	99US-0139763.
PR 21-JUN-1999;	99US-0139817.
PR 22-JUN-1999;	99US-0139899.
PR 23-JUN-1999;	99US-0140353.
PR 23-JUN-1999;	99US-0140354.
PR 24-JUN-1999;	99US-0140695.
PR 26-JUN-1999;	99US-0140623.
PR 29-JUN-1999;	99US-0140691.
PR 30-JUN-1999;	99US-0141287.
PR 01-JUL-1999;	99US-0141842.
PR 01-JUL-1999;	99US-0142154.
PR 02-JUL-1999;	99US-0142055.
PR 06-JUL-1999;	99US-0142390.
PR 08-JUL-1999;	99US-0142803.
PR 09-JUL-1999;	99US-0142920.
PR 12-JUL-1999;	99US-0143277.
PR 13-JUL-1999;	99US-0143542.
PR 15-JUL-1999;	99US-0143624.
PR 16-JUL-1999;	99US-0144005.
PR 16-JUL-1999;	99US-0144086.
PR 19-JUL-1999;	99US-0144325.
PR 19-JUL-1999;	99US-0144331.
PR 19-JUL-1999;	99US-0144332.
PR 19-JUL-1999;	99US-0144333.
PR 19-JUL-1999;	99US-0144334.
PR 19-JUL-1999;	99US-0144335.
PR 20-JUL-1999;	99US-0144352.
PR 20-JUL-1999;	99US-0144632.
PR 20-JUL-1999;	99US-0144884.
PR 21-JUL-1999;	99US-0144814.
PR 21-JUL-1999;	99US-0145086.
PR 21-JUL-1999;	99US-0145088.
PR 22-JUL-1999;	99US-0145085.

PR 22-JUL-1999;	99US-0145087.
PR 22-JUL-1999;	99US-0145089.
PR 22-JUL-1999;	99US-0145192.
PR 23-JUL-1999;	99US-0145145.
PR 23-JUL-1999;	99US-0145218.
PR 23-JUL-1999;	99US-0145224.
PR 26-JUL-1999;	99US-0145276.
PR 27-JUL-1999;	99US-0145513.
PR 27-JUL-1999;	99US-0145818.
PR 27-JUL-1999;	99US-0145919.
PR 28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 06-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 09-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 10-AUG-1999;	99US-0147935.
PR 11-AUG-1999;	99US-0148171.
PR 12-AUG-1999;	99US-0148319.
PR 13-AUG-1999;	99US-0148341.
PR 13-AUG-1999;	99US-0148565.
PR 16-AUG-1999;	99US-0148684.
PR 17-AUG-1999;	99US-0149368.
PR 18-AUG-1999;	99US-0149175.
PR 20-AUG-1999;	99US-0149426.
PR 20-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 23-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152363.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0158293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159329.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 41.0%; Score 168; DB 21; Length 369;
Best Local Similarity 99.4%; Pred.No. 2e-34;
Matches 168; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 gcgcatactgggcagaaaccttcctcaacattgcagagaagaaggctcccatctctgtg 71
DB 201 gcgcatactnggcagaaaccttcctcaacattgcagagaagaaggctcccatctctgtg 260
QY 72 tacaaacagacaacctccaaggttgacagagacgcgtgcagcggtgcacaagcgagaagaac 131
DB 261 tacaaacagacaacctccaaggttgacagagacgcgtgcagcggtgcacaagcgagaagaac 320
QY 132 ctcccgctacagcttcacatcacccgcgcctcttggaagttcattc 180
DB 321 ctcccgctacagcttcacatcacccgcgcctcttggaagttcattc 369

RESULT 8
AAT88035
ID AAT88035 standard; DNA; 1916 BP.
XX AC AAT88035;
XX
DT 22-DEC-1997 (first entry)
XX
DE Brevibacterium flavum 6-phosphogluconate dehydrogenase DNA.
XX
KW Brevibacterium flavum; 6-phosphogluconate dehydrogenase;
KW recombinant production; coryneform; bacterium; bacteria; ds.
XX
OS Brevibacterium flavum.
XX
FH Key
FT CDS Location/Qualifiers
FT 374..1852 /*tag= a
FT /product= 6-phosphogluconate_dehydrogenase
XX JF09224662-A.
XX

PD 02-SEP-1997.
XX
PF 23-FEB-1996; 96JP-0036346.
XX
PR 23-FEB-1996; 96JP-0036346.
XX
PA (MITU) MITSUBISHI CHEM CORP.
XX
DR WPI; 1997-484097/45.
DR P-PSDB; AAM27613.
XX
PT 6-phospho-beta-glucuronate dehydrogenase - can be recombinantly produced by
PT transforming corynebacterium bacteria with DNA molecule encoding it
XX
PS Claim 1; Pages 5-7; 8pp; Japanese.
XX
CC The present sequence encodes the Brevibacterium flavum JM-233
CC 6-phosphogluconate dehydrogenase (6-PD). The 6-PD can be
CC recombinantly produced by transforming corynebacterium bacteria with the
CC DNA molecule encoding it.
XX
SQ Sequence 1916 BP; 425 A; 596 C; 507 G; 388 T; 0 other;

Query Match 24.2%; Score 99.2; DB 18; Length 1916;
Best Local Similarity 54.9%; Pred. No. 2,4e-16;
Matches 219; Conservative 0; Mismatches 174; Indels 6; Gaps 1;

OY 12 ggcgtatcgtggcagaaaccttgcctcaacattgcagagaaaggtcccatctgtg 71
DB 443 gaagtaatgggtcacaacctcgcgcgaacttcgcgcgaacgagccacatctgcgtgc 502
OY 72 tacaagaagacaacactcgaaggtgagacgcgtgcagcgtgcgaagcagaagaac 131
DB 503 tacaacgcagcactgcgaacaacgcagcgtcgcgcgaacgcgtgcgaagcgaac 562
OY 132 ctcccgctcagcgttcacatgaacccgcgtccttgtagaagtcacatcagaacgag 191
DB 563 ttcattcccttcgcaacccgtcgaagag-----gtcgtagcattcccttggaagccagc 616
OY 192 gtggtatcatcgtcgtcgaagcgcgcgcaggttcgaagcagcagcagcagcagc 251
DB 617 cgcgcacatcatcgttcaagctgtgtaacgcacacgcagcagcagcagcagcagc 676
OY 252 gctcacttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 311
DB 677 gacgcacatgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 736
OY 312 gagaagagagagagagcagcagcagcagcagcagcagcagcagcagcagcagc 371
DB 737 attcgtcgcgagagagaaatctcgcagcagcagcagcagcagcagcagcagc 796
OY 372 ggagagaagagaggtgcccgaacgcccgtccttgatg 410
DB 797 ggcgcgagagagagcgcacatcagcagcagcagcagcagcagcagcagc 835

RESULT 9
AAH6570
ID AAH6570 standard; DNA; 1476 BP.
XX
AC AAH6570;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum coding sequence fragment SEQ ID NO: 1605.
XX
KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis; ds.
XX Corynebacterium glutamicum.
OS
PN EPI108790-A2.

XX 20-JUN-2001.
XX
PF 18-DEC-2000; 2000BP-0127688.
XX
PR 16-DEC-1999; 99JP-0377484.
XX
PR 07-APR-2000; 2000JP-0159162.
XX
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI; 2001-376931/40.
XX
DR P-PSDB; AAC91351.
XX
PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
XX
PS Claim 8; SEQ ID NO: 1605; 246bp + Sequence Listing; English.

XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Corynebacterium, and identifying a homologue of a gene derived
CC from Corynebacterium. Corynebacterium bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 1476 BP; 328 A; 489 C; 382 G; 277 T; 0 other;

Query Match 23.8%; Score 97.6; DB 22; Length 1476;
Best Local Similarity 54.6%; Pred. No. 5.9e-16;
Matches 218; Conservative 0; Mismatches 175; Indels 6; Gaps 1;

OY 12 ggcgtatcgtggcagaaaccttgcctcaacattgcagagaaaggtcccatctgtg 71
DB 70 gaagtaatgggtcacaacctcgcgcgaacttcgcgcgaacgagccacatctgcgtgtc 129
OY 72 tacaagaagacaacactcgaaggtgagacgcgtgcagcgtgcgaagcagaagaac 131
DB 130 tacaacgcagcactgcgaacaacgcagcgtcgcgcgaacgcgtgcgaagcgaac 189
OY 132 ctcccgctcagcgttcacatgaacccgcgtccttgtagaagtcacatcagaacgag 191
DB 190 ttcattcccttcgcaacccgtcgaagag-----gtcgtagcattcccttggaagccagc 243
OY 192 gtggtatcatcgtcgtcgaagcgcgcgcaggttcgaagcagcagcagcagcagc 251
DB 244 cgcgcacatcatcgttcaagctgtgtaacgcacacgcagcagcagcagcagcagc 303
OY 252 gctcacttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 311
DB 304 gatgcacatgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 363
OY 312 gagaagagagagagcagcagcagcagcagcagcagcagcagcagcagcagc 371
DB 364 attcgtcgcgagagagaaatctcgcagcagcagcagcagcagcagcagcagc 423
OY 372 ggagagaagagaggtgcccgaacgcccgtccttgatg 410
DB 424 ggcgcgagagagagcgcacatcagcagcagcagcagcagcagcagcagc 462


```
RESULT 10
AA71486
ID AAF71486 standard; DNA; 1537 BP.
XX
AC AAF71486;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:253.
XX
KW Corynebacterium glutamicum: carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polypeptide; enzyme;
XX diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.
XX
OS Corynebacterium glutamicum.
XX
FN W020100844-A2.
XX
PD 04-JAN-2001.
XX
PE 23-JUN-2000; 2000WO-IB00943.
XX
PR 25-JUN-1999; 9905-0141031.
PR 08-JUL-1999; 99DE-1031412.
PR 08-JUL-1999; 99DE-1031413.
PR 08-JUL-1999; 99DE-1031419.
PR 08-JUL-1999; 99DE-1031420.
PR 08-JUL-1999; 99DE-1031428.
PR 08-JUL-1999; 99DE-1031431.
PR 08-JUL-1999; 99DE-1031433.
PR 08-JUL-1999; 99DE-1031434.
PR 08-JUL-1999; 99DE-1031510.
PR 08-JUL-1999; 99DE-1031562.
PR 08-JUL-1999; 99DE-1031634.
PR 09-JUL-1999; 99DE-1032180.
PR 09-JUL-1999; 99DE-1032227.
PR 09-JUL-1999; 99DE-1032230.
PR 09-JUL-1999; 9905-0143208.
PR 14-JUL-1999; 99DE-1032924.
PR 14-JUL-1999; 99DE-1032973.
PR 14-JUL-1999; 99DE-1033003.
PR 27-AUG-1999; 99DE-1040765.
PR 31-AUG-1999; 9905-0151572.
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042123.
PR 03-SEP-1999; 99DE-1042125.
XX
PA (BAD1 ) BASF AG.
XX
PI Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
XX
DR WPI; 2001-061975/07.
XX
DR P-PsDB; AAB79369.
XX
PT New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT metabolism and oxidative phosphorylation protein for production or
PT modulation of production of fine chemicals e.g. amino acids,
XX carbohydrates or enzymes -
XX
PS Claim 3; Page 504-506; 1246pp; English.
XX
CC AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
```

```
CC metabolism and oxidative phosphorylation (SMP) proteins given in
CC AAB79243 to AAB 79633 which are involved in carbon metabolism and
CC energy production. The C. glutamicum SMP gene can be used in vectors
CC (II) for expression in host cells and production or modulation of
CC production of fine chemicals, such as, an organic acid, a proteinogenic
CC or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
CC a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
CC acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
CC cofactor, a polypeptide, or an enzyme. The presence of (I) or SMP proteins
CC (III) encoded by them are used for diagnosing the presence or activity of
CC Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
CC containing them are used to map genomes of organisms related to
CC C. glutamicum, identify and localize C. glutamicum sequences of interest,
CC in evolutionary studies, in determining SMP protein regions required
CC for function, in modulating SMP protein activity, in modulating the
CC metabolism of sugars, and in modulating high-energy molecule production
CC in a cell (i.e. ATP, NADPH).
XX
SQ Sequence 1537 BP; 341 A; 512 C; 392 G; 292 T; 0 other:
XX
Query Match 23.8%; Score 97.6; DB 22; Length 1537;
Best Local Similarity 54.6%; Pred. No. 6e-16;
Matches 218; Conservative 0; Mismatches 175; Indels 6; Gaps 1;
QY 12 gcgcgtacgggcaagaacctgcccacatctgcagagaagggtcccatctctgtg 71
Db 146 gcagtaatggtctcaaacctcgcgcgaacttgcgcgaacgcaacactgtcgtctc 205
QY 72 tacacaggaacacctcaaggttgagcagacgctcagcgttgccaaaggcagaagaac 131
Db 206 tacacacgcgcacatgcacaaacagcaagctcgcgcatcagcgtccgcgaaggcaac 265
QY 132 ctccgcgtcagcgtctccatcagccccgcgtctcttgtaagtcattcagaagcagcg 191
Db 266 ttaatccctctcgaacgctcgaga-----gtcgttcgacatcccttgaaagccacgc 319
QY 192 gtgtgatactctcgtcgaagcgcgcgcgttgtagcagaacatcgcgacgtctgca 251
Db 320 cgcgcacatcactatgttcaagtgtaagccacgcagcagtcactcaacacgcgtgca 379
QY 252 gctcacttgagcagagcgactcactcatcgatggggggaacgagtgtgacagaacagc 311
Db 380 gatgcacatgcagaaagcgacatcactcagcagcggaacccctcaccacgcgaccc 439
QY 312 gaagaggaaggaagggccatgagagcgcgccctcgtatcttgatctggatgggtctct 371
Db 440 attcgtcgcgagaagaatctccgacgcggtctccactcgtcgtgctgtatctcc 499
QY 372 ggaagaaaggaagggtgcccgcgaagcgcgtctctgagt 410
Db 500 ggcgcgcaagaagcgcatcacaagcgcacatcatgt 538
RESULT 11
AA71485
ID AAF71485 standard; DNA; 1575 BP.
XX
AC AAF71485;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:251.
XX
KW Corynebacterium glutamicum: carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polypeptide; enzyme;
XX diagnosis; Corynebacterium diphtheriae; evolutionary study; ds.
XX
OS Corynebacterium glutamicum.
```


XX WO200100844-A2.
 XX 04-JAN-2001.
 PD
 XX
 PE 23-JUN-2000; 2000MO-IB00943.
 XX
 PR 25-JUN-1999; 99US-0141031.
 PR 08-JUL-1999; 99DE-1031412.
 PR 08-JUL-1999; 99DE-1031413.
 PR 08-JUL-1999; 99DE-1031419.
 PR 08-JUL-1999; 99DE-1031420.
 PR 08-JUL-1999; 99DE-1031424.
 PR 08-JUL-1999; 99DE-1031428.
 PR 08-JUL-1999; 99DE-1031431.
 PR 08-JUL-1999; 99DE-1031433.
 PR 08-JUL-1999; 99DE-1031434.
 PR 08-JUL-1999; 99DE-1031510.
 PR 08-JUL-1999; 99DE-1031562.
 PR 08-JUL-1999; 99DE-1031634.
 PR 09-JUL-1999; 99DE-1032180.
 PR 09-JUL-1999; 99DE-1032227.
 PR 09-JUL-1999; 99DE-1032230.
 PR 09-JUL-1999; 99US-0143208.
 PR 14-JUL-1999; 99DE-1032824.
 PR 14-JUL-1999; 99DE-1032873.
 PR 14-JUL-1999; 99DE-1033005.
 PR 27-AUG-1999; 99DE-1040765.
 PR 31-AUG-1999; 99US-0151572.
 PR 03-SEP-1999; 99DE-1042076.
 PR 03-SEP-1999; 99DE-1042079.
 PR 03-SEP-1999; 99DE-1042086.
 PR 03-SEP-1999; 99DE-1042087.
 PR 03-SEP-1999; 99DE-1042088.
 PR 03-SEP-1999; 99DE-1042095.
 PR 03-SEP-1999; 99DE-1042123.
 PR 03-SEP-1999; 99DE-1042125.
 XX
 PA (BAD) BASF AG.
 PI
 XX
 DR Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
 XX
 DR WPI: 2001-061975/07.
 XX
 PT P-PSDB; AAB79368.
 PT
 PT New isolated *Corynebacterium glutamicum* nucleic acid encoding a sugar
 PT metabolism and oxidative phosphorylation protein for production or
 PT modulation of production of fine chemicals e.g. amino acids,
 PT carbohydrates or enzymes -
 XX
 PS Claim 3; Page 500-502; 1246pp; English.
 XX
 CC AAF71360 to AAF71750 encode the *Corynebacterium glutamicum* sugar
 CC metabolism and oxidative phosphorylation (SMP) proteins given in
 CC AAB79243 to AAB 79633 which are involved in carbon metabolism and
 CC energy production. The C. glutamicum SMP gene can be used in vectors
 CC (II) for expression in host cells and production or modulation of
 CC production of fine chemicals, such as, an organic acid, a proteinogenic
 CC or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
 CC a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
 CC acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
 CC cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
 CC (III) encoded by them are used for diagnosing the presence or activity of
 CC *Corynebacterium diphtheriae* in a subject. (I), (II), (III) or host cells
 CC containing them are used to map genomes of organisms related to
 CC C. glutamicum, identify and localise C. glutamicum sequences of interest,
 CC in evolutionary studies, in determining SMP protein regions required
 CC for function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. ATP, NADPH).
 CC
 XX Sequence 1575 BP; 352 A; 521 C; 399 G; 303 T; 0 other;

Query Match 23.8%; Score 97.6; DB 22; Length 1575;
 Best Local Similarity 54.6%; Pred. No. 6e-16;
 Matches 218; Conservative 0; Mismatches 175; Indels 6; Gaps 1;
 QY 12 gcggtcatgggagcaacctgcccctcaacatgacagagaagaaggtcccatctctg 71
 Db 146 gcagtaatgggctcaaacctcgcccgcaacttcgcccgaagcaacatgctgc 205
 QY 72 tacacagagacaacctcaaggggtgagagacggtgcaggtgcgaaggaaggaac 131
 Db 206 tacacccgacacatgacaacacgacaagctcatcgccagtcggtccgaaggaac 265
 QY 132 ctcccgctcagcgctccatccatccaccgctccctttgtgaatccattcagaagccacg 191
 Db 266 ttatccctcttgcaacgcgtcgaaga-----gtctgtagcatccctggaaagccacgc 319
 QY 192 gtggtatcatgctcgtcaagggccgagccagttgacacagacatcgacgtcgca 251
 Db 320 cgcgcctcatcatgctcaggtgctgtaacgcccacgacgtcatcaacacagctgca 379
 QY 252 gtcacttgagcgagcgagctgcatcatcgatgggggagacgagtggtacgagaacag 311
 Db 380 gatgcataggaggaagcgacatcatcgcagcgagcaagccctctacacgacaac 439
 QY 312 gagagagaggaagggcattggaagagcgcgccctnctgtaattcgtgcatgggtctct 371
 Db 440 atctgcgcgagaagaatctccgcagcggtctccattcgtcggtgctggtatctcc 499
 QY 372 gagaggaaggaaggtgcccgaagcgccgctctctatg 410
 Db 500 ggcgcgagaagagcgacctaacgcccatacatcatg 538
 RESULT 12
 AAI65505
 ID AAI65505 standard; DNA; 2335 BP.
 AC
 XX
 AC AAI65505;
 XX
 DT 10-DEC-2001 (first entry)
 XX
 DE Nucleotide sequence of the *Corynebacterium* gnd gene.
 XX
 XX gnd gene; 6-phosphogluconate dehydrogenase; L-amino acid; L-lysine;
 KM L-threonine; L-tryptophan; ss.
 XX
 OS *Corynebacterium glutamicum*.
 XX
 XX Key Location/Qualifiers
 FH 474..1853
 FT CDS /*tag= a
 FT /product= "6-phosphogluconate dehydrogenase"
 XX
 PN WO200171012-A1.
 XX
 PD 27-SEP-2001.
 XX
 PF 05-JUL-2000; 2000MO-EP06299.
 XX
 PR 20-MAR-2000; 2000US-0531265.
 XX
 XX (DEGS) DEGUSSA AG.
 PA (UYNA-) UNIV NAT IRELAND.
 XX
 PI Dunltau IK, McCormack A, Stapellon C, Burke K, Moeckel B;
 DR WPI: 2001-602792/68.
 DR P-PSDB; AAG79095.
 XX
 PT Preparing L-amino acids by fermenting *coryneform* bacteria transformed
 PT with the 6-phosphogluconate dehydrogenase gene is particularly useful
 PT to produce L-lysine and L-threonine -


```
XX XX Example 2; Page 43-46; 59pp; English.
PS CC The present sequence represents the gnd gene of Corynebacterium
XX CC glutamicum. The gnd gene encodes a 6-phosphogluconate dehydrogenase
XX CC enzyme. The gnd gene is used to produce recombinant bacteria for use
XX CC in the course of the invention. The specification describes a method
XX CC for the fermentative preparation of L-amino acids, in particular
XX CC L-lysine, L-threonine, and L-tryptophan. The method comprises fermenting
XX CC coryneform bacteria in which the gnd gene is amplified and overexpressed.
XX CC The L-amino acids produced are used in animal nutrition, human medicine
XX CC and the pharmaceuticals industry.
SQ Sequence 2335 BP; 533 A; 695 C; 606 G; 501 T; 0 other;

Query Match          23.8%; Score 97.6; DB 22; Length 2335;
Best Local Similarity 54.6%; Pred. No. 6.7e-16;
Matches 218; Conservative 0; Mismatches 175; Indels 6; Gaps 1;

QY 12 ggcgtcatgaggcagaaccttcgacctcaacattgcagagaagggtcccatctctgtg 71
DB 543 gcaatgatggtcacaacctgcgcgcgaaccttcgcgcgaagcaacacatgtgcctgc 602
QY 72 taacaagagacaacctccaaggtgagcagcagcgtgcagcagcagcagcagcagcagc 131
DB 603 taacaacgcagacacgcagacaacccgacaagctcaccgcagcagcagcagcagcagc 662
QY 133 ctcccgcttaagcgttcacaccccgctccctgttggaagttccatcagaagcagcag 191
DB 663 ttcatccctctgcacacgcgtcgaga-----gttcgttagacatcccttggaagcagcag 716
QY 192 gtggtgcatcagctcgtcaaaagcgcgcgcgcaggttgacacacacacacacacacac 251
DB 717 cgcgcacacacacacacacacacacacacacacacacacacacacacacacacacacac 776
QY 252 gctcacttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 311
DB 777 gatgcacatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 836
QY 312 gaggaggaaggagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 371
DB 837 attcgtcgcagagaggaatctccgcacgcgctcaccatctgcgtgctgctatctcc 896
QY 372 ggaagaaaggaggtgcccgcagcagcagcagcagcagcagcagcagcagcagcagc 410
DB 897 ggcgcgcagagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 935

RESULT 13
AAH68528/c
ID AAH68528 standard; DNA; 349980 BP.
XX
XX AAH68528;
AC
XX 26-SEP-2001 (first entry)
DE C glutamicum coding sequence fragment SEQ ID NO: 7063.
XX
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KM organic acid synthesis; ds.
XX
XX Corynebacterium glutamicum.
OS
XX BP1108790-A2.
PN
XX
XX 20-JUN-2001.
PD
XX
XX 18-DEC-2000; 2000EP-0127688.
PF
XX
XX 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
```

```
XX XX (KYOWA ) KYOWA HAKKO KOGYO KK.
PA
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX PI Tetsushi N, Senoh A, Ikeda M, Ozaki A;
XX PI
XX WPI: 2001-376931/40.
DR
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
XX PT expression point of a gene, measuring expression of a gene, analysing
XX PT expression profile or pattern of a gene and identifying homologous gene
XX PT
XX
XX
PS Disclosure; SEQ ID NO: 7063; 246pp + Sequence Listing; English.
XX
XX
```

```
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
XX
```

SQ Sequence 349980 BP; 80727 A; 91049 C; 97618 G; 80586 T; 0 other;

```
Query Match          23.8%; Score 97.6; DB 22; Length 349980;
Best Local Similarity 54.6%; Pred. No. 2.4e-15;
Matches 218; Conservative 0; Mismatches 175; Indels 6; Gaps 1;

QY 12 ggcgtcatgaggcagaaccttcgacctcaacattgcagagaagggtcccatctctgtg 71
DB 331747 GCAGTAAATGGGCTCAAAACCTCGCGCAACTTCGCCCGCAAGCGAACACTGCTCTCTC 331688
QY 72 taacaagagacaacctccaaggtgagcagcagcagcagcagcagcagcagcagcagcagc 131
DB 331687 TACAAACCGACGACCTGACAAAACCGACAAAGCTACATCGCGATCAGCGCTCCGAAGCGAAC 331628
QY 133 ctcccgcttaagcgttcacaccccgctccctgttggaagttccatcagaagcagcag 191
DB 331627 TTCATCCCTTCAGCAACCGTCAAGA-----GTTCTGAAGCATCCCTCGAAAAGCGACGC 331574
QY 192 gtggtgcatcagctcgtcaaaagcgcgcgcgcaggttgacacacacacacacacacac 251
DB 331573 CGCGCCATCATCATGTGTTTCAGGCTGTGAACGCCACGACGAGTCATCAACCACTGGCA 331514
QY 252 gctcacttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 311
DB 331513 GATGCCATGGAAGAGGACATCATCATCGACGGCGGCAAGCGCCCTTACACACCGACAC 331454
QY 312 gaggaggaaggagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 371
DB 331453 ATTGCTCGGAGAAAGAAATCTCCGACAGCGGTCTCCACTTCGTGCTGCTGATCTCC 331394
QY 372 ggaagaaaggaggtgcccgcagcagcagcagcagcagcagcagcagcagcagcagc 410
DB 331393 GCGCGGCAAGAAAGCGGCACTCAACGCGCCATCATCATG 331355

RESULT 14
AAH68529/c
ID AAH68529 standard; DNA; 349980 BP.
XX
XX AAH68529;
AC
XX 26-SEP-2001 (first entry)
DE
XX
```



```
Db 213 accaaagtcgattggagccgactcgctcgagagacatggtctccaagctgaagagcccccgg 272
QY 192 gtgtgatcatgctcgtcaaggccggcgccagttgaaccaacatcgcgacgctcgca 251
Db 273 aaggtcatgctgtgtcaaggcttgaaagtgcagtcgacgacttcattccagcagctggtg 332
QY 252 gtcactggagcaggcgactgcattcgatgggggaaagtggtacgagaacag 311
Db 333 ccgctgcttcgcgcgcgcgcgcatgcatcgcattggtgcaactcggagttacaggaacaca 392
QY 312 gagaggaaggagaaagccattgagagagcgccctnctgtattcttggcatgtgtcttct 371
Db 393 tctcgccgctgtagcagtagtccaaacttgccctgctcttcgtcgtgattccgcgtgagc 452
QY 372 gaggaaggaagggtgcccgcacagggccgctccttgatg 410
Db 453 ggtggcgaaggagggcgccgcacaggaacctcgctgagt 491
```

Search completed: July 3, 2002, 07:56:21
Job time: 24336 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 07:41:01 ; Search time 185.98 Seconds
(without alignments)
541,508 Million cell updates/sec

Title: US-09-300-482-14

Perfect score: 410
Sequence: 1 cccacgcgtccgcgcatg.....gcacgcgcgtccttgatg 410

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCITUS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56.4	13.8	2087	4 US-09-058-692-1	Sequence 1, Appl
2	56.4	13.8	2087	4 US-09-584-628-1	Sequence 1, Appl
3	37.4	9.1	1524	2 US-08-716-942-24	Sequence 24, Appl
4	35.6	8.7	11958	4 US-09-134-246-8	Sequence 8, Appl
5	35.6	8.7	939	4 US-09-105-390-45	Sequence 45, Appl
6	35.6	8.7	1020	4 US-09-105-390-61	Sequence 61, Appl
7	35.6	8.7	2612	4 US-09-105-390-7	Sequence 7, Appl
8	35.6	8.7	8051	2 US-08-576-626A-2	Sequence 2, Appl
9	35.2	8.6	2990	1 US-08-572-951-1	Sequence 1, Appl
10	35.2	8.6	3476	3 US-08-630-916A-47	Sequence 47, Appl
11	35	8.5	2830	2 US-09-010-928B-1	Sequence 1, Appl
12	34.4	8.4	2277	1 US-08-676-967-5	Sequence 5, Appl
13	34.4	8.4	2277	1 US-08-676-974-5	Sequence 5, Appl
14	34.4	8.4	2277	2 US-09-098-487-5	Sequence 13, Appl
15	34	8.3	2713	3 US-08-804-439A-13	Sequence 13, Appl
16	34	8.3	2713	3 US-08-720-229-13	Sequence 13, Appl
17	34	8.3	3472	6 5244792-2	Patent No. 5244792
18	34	8.3	5020	3 US-08-938-291A-3	Sequence 3, Appl
19	33.8	8.2	1910	5 PCT-US92-05532-1	Sequence 1, Appl
20	33.8	8.2	2261	1 US-08-272-882D-1	Sequence 1, Appl
21	33.8	8.2	7218	1 US-08-232-463-14	Sequence 14, Appl
22	33.2	8.1	2040	2 US-08-533-669A-5	Sequence 5, Appl
23	33	8.0	1586	4 US-09-392-772-3	Sequence 3, Appl
24	33	8.0	49272	1 US-08-614-770A-1	Sequence 1, Appl
25	32.8	8.0	1312	1 US-08-205-506A-1	Sequence 1, Appl
26	32.8	8.0	1312	5 PCT-US94-02389-1	Sequence 1, Appl
27	32.8	8.0	3318	4 US-09-593-589-3	Sequence 3, Appl

28	32.4	7.9	28804	2 US-08-592-874-1	Sequence 1, Appl
29	32.4	7.9	28804	3 US-09-096-942-2	Sequence 2, Appl
30	32.4	7.9	28804	3 US-09-096-867-2	Sequence 2, Appl
31	32.2	7.9	720	4 US-09-459-956-1	Sequence 1, Appl
32	32.2	7.9	1929	2 US-08-818-253-1	Sequence 1, Appl
33	32.2	7.9	1929	2 US-08-818-253-5	Sequence 5, Appl
34	32.2	7.9	1929	4 US-08-818-252-1	Sequence 1, Appl
35	32.2	7.9	1929	4 US-08-818-252-5	Sequence 5, Appl
36	32.2	7.9	1959	2 US-08-818-252-3	Sequence 3, Appl
37	32.2	7.9	1959	4 US-08-818-252-3	Sequence 3, Appl
38	32.2	7.9	1971	2 US-08-818-252-7	Sequence 7, Appl
39	32.2	7.9	1971	4 US-08-818-252-7	Sequence 7, Appl
40	32.2	7.9	2680	2 US-08-533-306A-5	Sequence 5, Appl
41	32.2	7.9	2680	2 US-08-742-923A-5	Sequence 5, Appl
42	32.2	7.9	2887	2 US-08-533-306A-3	Sequence 3, Appl
43	32.2	7.9	2887	2 US-08-742-923A-3	Sequence 3, Appl
44	32	7.8	720	3 US-09-094-359-5	Sequence 5, Appl
45	32	7.8	720	3 US-09-094-359-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-09-058-692-1
; Sequence 1, Application US/09058692
; Patent No. 6162618
; GENERAL INFORMATION:
; APPLICANT: Warren, Richard L.
; TITLE OF INVENTION: 6-phosphogluconate dehydr
; TITLE OF INVENTION: 4
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/058,692
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10140
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2087 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-058-692-1

Query Match 13.8%; Score 56.4; DB 4; Length 2087;
Best Local Similarity 48.4%; Pred. No. 3.3e-06;
Matches 193; Conservative 0; Mismatches 197; Indels 9; Gaps 1;
QY 12 gcgcatggggcagaacctgcctcaacatgcagagaaggtcccatctctgtg 71


```

Db 44 GCGGTAATGGGCTGTAACCTTGCCCTTAATATGAAATCWCAGTGGTTACACAGTTGCTATC 103
Qy 72 tacacagagacaactccaaggctgagacagacgcgcgcgttcgaagccgaagaaac 131
Db 104 TACAACCTGATGAAAGAAAAGGAGATGATGATGCTTCCATCCGAAAGAACTTT 163
Qy 132 ctcccgctcagcgtccatcagcccgctcctcttgtaagtcacatcagaagccagc 191
Db 164 GTACCAAGCT-----ATGACGTTGAAGTTTGTAAACTCAATCGAAAACCTGCT 214
Qy 192 gtgagatcagctcgtcgaagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 251
Db 215 CGTATCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 274
Qy 252 gtcacacttgagacagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 311
Db 275 CCACACCTTGACAAAGGATGATGATGATGATGATGATGATGATGATGATGATG 334
Qy 312 gagagagagagagagagagagagagagagagagagagagagagagagagagag 371
Db 335 ATCCGTCGTAATGAAGATTTGGCAAACTCWCAGTATCAACTTTATCGGTAAGT 394
Qy 372 gagagagagagagagagagagagagagagagagagagagagagagagagagag 410
Db 395 GGTGTGAAAAAGGTGCCCTTGAAGGTCTTCTATCATG 433

```

```

RESULT 2
US-09-584-628-1
; Sequence 1, Application us/09584628
; Patent No. 6309866
; GENERAL INFORMATION:
; APPLICANT: Warren, Richard L.
; TITLE OF INVENTION: 6-phosphogluconate dehydr
; TITLE OF INVENTION: ogenase
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Decher, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; City: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/584,628
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/058,692
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10140
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2087 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-584-628-1

```

```

Query Match 13.8% Score 56.4; DB 4; Length 2087;
Best Local Similarity 48.4%; Pred. No. 3..3e-06;
Matches 193; Conservative 0; Mismatches 197; Indels 9; Gaps 1;
Qy 12 gcggtcacttgagacagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 71
Db 44 GCGGTAATGGGCTGTAACCTTGCCCTTAATATGAAATCWCAGTGGTTACACAGTTGCTATC 103
Qy 72 tacacagagacaactccaaggctgagacagacgcgcgcgttcgaagccgaagaaac 131
Db 104 TACAACCTGATGAAAGAAAAGGAGATGATGATGATGATGATGATGATGATGATGATG 163
Qy 132 ctcccgctcagcgtccatcagcccgctcctcttgtaagtcacatcagaagccagc 191
Db 164 GTACCAAGCT-----ATGACGTTGAAGTTTGTAAACTCAATCGAAAACCTGCT 214
Qy 192 gtgagatcagctcgtcgaagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 251
Db 215 CGTATCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 274
Qy 252 gtcacacttgagacagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 311
Db 275 CCACACCTTGACAAAGGATGATGATGATGATGATGATGATGATGATGATGATG 334
Qy 312 gagagagagagagagagagagagagagagagagagagagagagagagagagag 371
Db 335 ATCCGTCGTAATGAAGATTTGGCAAACTCWCAGTATCAACTTTATCGGTAAGT 394
Qy 372 gagagagagagagagagagagagagagagagagagagagagagagagagagag 410
Db 395 GGTGTGAAAAAGGTGCCCTTGAAGGTCTTCTATCATG 433

```

```

RESULT 3
US-08-716-942-24
; Sequence 24, Application US/08716942
; Patent No. 5849491
; GENERAL INFORMATION:
; APPLICANT: Terragen Diverslty Inc.
; APPLICANT: Radomski, Christopher C. A.
; APPLICANT: Seow, Kah Tong
; APPLICANT: Warren, R. Anthony J.
; APPLICANT: Vap, Wal Ho
; TITLE OF INVENTION: METHOD FOR ISOLATING XYLANASE GENE
; TITLE OF INVENTION: SEQUENCES FROM SOIL DNA, COMPOSITIONS USEFUL IN SUCH METHOD
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Opedahl & Larson
; STREET: 1992 Commerce Street, Suite 309
; CITY: Yorktown Heights
; STATE: NY
; COUNTRY: USA
; ZIP: 10598-4412
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS 5.0
; SOFTWARE: Word Perfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/716,942
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/004,157
; FILING DATE: 20-Sep-95
; ATTORNEY/AGENT INFORMATION:
; NAME: Marina T. Larson
; REGISTRATION NUMBER: 32,038
; REFERENCE/DOCKET NUMBER: TERR-P-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 245-3252
; TELEFAX: (914) 962-4330

```


1

```
TELEX:
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1524
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; ORGANISM:
; FEATURE:
; NAME/KEY: sequence of xylanase gene identified by
; NAME/KEY: amplification of xylanase fragments from soil
US-08-716-942-24
```

```
Query Match 9.1%; Score 37.4; DB 2; Length 1524;
Best Local Similarity 50.6%; Pred. No. 0.34;
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
```

```
QY 189 cgggtggtatcatctctcgtcaaggcgcgccagttgaccagacatcggaagctc 248
DB 595 CCGCGAGTGGGAGCGGGTACATCGCCGACGCGCTCGCAGCCGCGACCGCCGACCC 654
QY 249 ggaagtcacttgagagcaagcgactgcatcagatggggagcaagtggtacgaagac 308
DB 655 AGTGGCCAGCTGTACATCAACAGCATCAGAGCGGAGGAGCCCAAGAACAGCC 714
QY 309 acggaagagagagagagagcgatggagagcgcgctcgtatcttgcatggg 364
DB 715 ATGTACAGCCTGTGTGACATCCTGCTGCGCGAGGGGTGCCGATCAACGCGGTGG 770
```

RESULT 4

```
US-09-134-246-8
; Sequence 8, Application US/09134246B
; Patent No. 6207377
; GENERAL INFORMATION:
; APPLICANT: Wayne, Jay
; TITLE OF INVENTION: Method For Construction Of Thermus-E. coli Shuttle
; TITLE OF INVENTION: Vectors And Identification Of Two Thermus Plasmid
; FILE REFERENCE: Thermus Shuttle Vector
; CURRENT APPLICATION NUMBER: US/09/134,246B
; CURRENT FILING DATE: 1998-08-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 8
; LENGTH: 11958
; TYPE: DNA
; ORGANISM: Thermus sp.
US-09-134-246-8
```

```
Query Match 8.7%; Score 35.8; DB 4; Length 11958;
Best Local Similarity 46.8%; Pred. No. 1.7;
Matches 152; Conservative 0; Mismatches 167; Indels 6; Gaps 1;
```

```
QY 9 tcgcggtcatatgggagcaaccttgccctcaacattgcagagaaaggttcccatctct 68
DB 7959 tccataggaacaagagggatgtatgccccacccttaagtagtaacacctcaaccgcgc 8018
QY 69 gtgta-----caacagaaacacctcccaagtgtgaagagaaacggtgacgtccaaagca 122
DB 8019 gtgaagagagggccaccttgaagcagctacataaggtcattcttcgtggcggaagagcc 8078
QY 123 gaaggaacatccctccgcttaagcgttcacgcccgcgtctcttggtaagcattcag 182
DB 8079 caaaagacctggtctctactctcaaggagcccgaggtccctcaaggtgtgagagagctctg 8138
```

```
QY 183 aagcacgggtgtgtatcatctctgtcaaggccgcgcgcccagttgaccagaccatcgcg 242
DB 8139 aagcccttggtcctgtgaaggtctccacaggggaggtgtcgagcagagagggaaacccg 8198
QY 243 agcgtgcagcttacttgtagcagagcgagctgcatatctgaggggagagagtgtagc 302
DB 8199 atgggggacctgacctgagagagatcatgccttcacagcgccgggtgaagagttcaag 8258
QY 303 gagaacacgagagagagaggaag 327
DB 8259 gagatgagaggaagagttcacgaag 8283
```

RESULT 5

```
US-09-105-390-45/c
; Sequence 45, Application US/09105390
; Patent No. 6288303
```

```
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Raymond
; TITLE OF INVENTION: Rice Beta-Glucanase Enzymes
; TITLE OF INVENTION: and Genes
; NUMBER OF SEQUENCES: 66
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
```

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/105,390

FILING DATE: Filed herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/050,675

FILING DATE: 25-JUN-97

ATTORNEY/AGENT INFORMATION:

NAME: Bethory, Joanne R.

REGISTRATION NUMBER: P42,995

REFERENCE/DOCKET NUMBER: 2000-0455.30

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-324-0880

TELEFAX: 650-324-0960

TELEX:

INFORMATION FOR SEQ ID NO: 45:

SEQUENCE CHARACTERISTICS:

LENGTH: 939 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 1..939

OTHER INFORMATION:

US-09-105-390-45

```
Query Match 8.7%; Score 35.6; DB 4; Length 939;
Best Local Similarity 45.9%; Pred. No. 0.88;
Matches 119; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
```

```
QY 92 ggtgaagagagacgtgtgcaaggtgcaagagaaaccttccgttcaaggttcca 151
DB 621 GGTGATGTGACGCCGCGTGTGTCACGACGCCGACTGCAACCCGACGACGACGCGTA 562
QY 152 tgaccgggtctcttgttaagttcatcagaagcaggggtgtgtagtatctgtctca 211
```


Db 561 CCCGAGCTGCACGCTGCTCGGGTCCGCCGCTACGCAAAAGTAGGGGTACACGTTACACG 502
QY 212 ggcgcgcgcgcagcagttgacacagccatcgcgcagctgcagctcaacttgagcagcgca 271
Db 501 CAGCGCGCGCCCTGTGACACGGAGTAGGCAACATCGGCCACTGTACGGCGACGCGCC 442
QY 272 ctgcatactgatatgggggagagagtggtacagaaacacgagagagagagagagagccat 331
Db 441 CTCGAGAACGCGCCCTCGACGCGCGGTACGACGCGGACGACGCGGACGCTGGCGACGCGC 382
QY 332 ggaagagcgcgccctnctg 350
Db 381 GGTGCTGACCGGATGCCG 363

RESULT 6
US-09-105-390-61/c

; Sequence 61, Application US/09105390
; Patent No. 6288303
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Raymond
; TITLE OF INVENTION: Rice Beta-Glucanase Enzymes
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,390
; FILING DATE: Filed herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/050,675
; FILING DATE: 25-JUN-97
; ATTORNEY/AGENT INFORMATION:
; NAME: Pelthory, Joanne R.
; REGISTRATION NUMBER: P42,995
; REFERENCE/DOCKET NUMBER: 2000-0455.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; TELEX:
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1020 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1..1020
; OTHER INFORMATION:
; US-09-105-390-61

Query Match 8.7%; Score 35.6; DB 4; Length 1020;
Best Local Similarity 45.9%; Pred. No. 0.9; Mismatches 140; Indels 0; Gaps 0;

Db 92 ggtgacgagacgctgacgagctgccaagcagaagaacctcccgctacagcttcca 151
QY 702 GGTGTATGTACAGCGCGCTGCTACCGACGCGGACGCTGCGACCGCGCGCTA 643
Db 152 tgacccgcgcctcttggaagtcattcagaagcagcggtggtgacatcgtctgtaaa 211

Db 642 CCCGAGCTGCACCTCTCTCGGTCCGCCGCTACGCAAAAGTAGGGGTACACGTTACACG 583
QY 212 ggcgcgcgcgcagttgacacagccatcgcgcagctgcagctcaacttgagcagcgca 271
Db 582 CAGCGCGCGCCCTGTGACACGGAGTAGGCAACATCGGCCACTGTACGGCGACGCGCC 523
QY 272 ctgcatactgatatgggggagagagtggtacagaaacacgagagagagagagagccat 331
Db 522 CTCGAGAACGCGCCCTCGACGCGCGGTACGACGCGGACGACGCGGACGCTGGCGACGCGC 463
QY 332 ggaagagcgcgccctnctg 350
Db 462 GGTGCTGACCGGATGCCG 444

RESULT 7
US-09-105-390-7/c

; Sequence 7, Application US/09105390
; Patent No. 6288303
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Raymond
; TITLE OF INVENTION: Rice Beta-Glucanase Enzymes
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,390
; FILING DATE: Filed herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/050,675
; FILING DATE: 25-JUN-97
; ATTORNEY/AGENT INFORMATION:
; NAME: Pelthory, Joanne R.
; REGISTRATION NUMBER: P42,995
; REFERENCE/DOCKET NUMBER: 2000-0455.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2612 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-105-390-7

Query Match 8.7%; Score 35.6; DB 4; Length 2612;
Best Local Similarity 45.9%; Pred. No. 1.2;
Matches 119; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Db 92 ggtgacgagacgctgacgagctgccaagcagaagaacctcccgctacagcttcca 151
QY 2294 GGTGTATGTACAGCGCGCTGCTACCGACGCGGACGCTGCGACCGCGCGCTA 2235
Db 152 tgacccgcgcctcttggaagtcattcagaagcagcggtggtgacatcgtctgtaaa 211
Db 2234 CCCGAGCTGCACCTCTCTCGGTCCGCCGCTACGCAAAAGTAGGGGTACACGTTACACG 2175

Db	721	CGCACCGGCGGCGACGACGCGCTGAGACCTCCGCGGGGCGGCCCGTCCGCCACGGCGG	662
Qy	218	CGCGCGAGTTGACCAAGACCAATCGAGAGCTCGCAGCTCACTTGGAGCAAGGCGACTGAT	277
Db	661	CGGCGCAAGCGCCCGCGCGAGAGAGAGGACACGCGCGCCGCCACGCGCCGAGGCGCTCCGCAT	602
Qy	278	CATCGATCGGAGGAGAAC	293
Db	601	TGCGGAGAGGAGAGC	586

```

1 RESULT 10
2 US-08-630-916A-47
3 ; Sequence 47, Application US/08630916A
4 ; Patent No. 6011137
5 ;
6 ; GENERAL INFORMATION:
7 ; APPLICANT: Pirozzi, Gregorio
8 ; APPLICANT: Kay, Brian K.
9 ; APPLICANT: Fowlkes, Dana M.
10 ; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
11 ; TITLE OF INVENTION: POLYPEPTIDES HAVING MW DOMAINS AND METHODS OF USING SAME
12 ; NUMBER OF SEQUENCES: 124
13 ; CORRESPONDENCE ADDRESS:
14 ; ADDRESSEE: Pennie & Edmonds
15 ; STREET: 1155 Avenue of the Americas
16 ; CITY: New York
17 ; STATE: New York
18 ; COUNTRY: United States
19 ; ZIP: 10036-2711
20 ;
21 ; COMPUTER READABLE FORM:
22 ; MEDIUM TYPE: Floppy disk
23 ; COMPUTER: IBM PC compatible
24 ; OPERATING SYSTEM: PC-DOS/MS-DOS
25 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
26 ;
27 ; CURRENT APPLICATION DATA:
28 ; APPLICATION NUMBER: US/08/630,916A
29 ; FILING DATE: 03-APR-1996
30 ; CLASSIFICATION: 435
31 ; ATTORNEY/AGENT INFORMATION:
32 ; NAME: MISROCK, S. LESLIE
33 ; REGISTRATION NUMBER: 18,872
34 ; REFERENCE/DOCKET NUMBER: 1101-203
35 ; TELECOMMUNICATION INFORMATION:
36 ; TELEPHONE: (212) 790-9090
37 ; TELEFAX: (212) 896-8864/9741
38 ; INFORMATION FOR SEQ ID NO: 47:
39 ; SEQUENCE CHARACTERISTICS:
40 ; LENGTH: 3476 bases
41 ; TYPE: nucleic acid
42 ; STRANDEDNESS: single
43 ; TOPOLOGY: linear
44 ;
45 ; MOLECULE TYPE: DNA
46 ;
47 ;
48 ;
49 ;
50 ;
51 ;
52 ;
53 ;
54 ;
55 ;
56 ;
57 ;
58 ;
59 ;
60 ;
61 ;
62 ;
63 ;
64 ;
65 ;
66 ;
67 ;
68 ;
69 ;
70 ;
71 ;
72 ;
73 ;
74 ;
75 ;
76 ;
77 ;
78 ;
79 ;
80 ;
81 ;
82 ;
83 ;
84 ;
85 ;
86 ;
87 ;
88 ;
89 ;
90 ;
91 ;
92 ;
93 ;
94 ;
95 ;
96 ;
97 ;
98 ;
99 ;
100 ;
101 ;
102 ;
103 ;
104 ;
105 ;
106 ;
107 ;
108 ;
109 ;
110 ;
111 ;
112 ;
113 ;
114 ;
115 ;
116 ;
117 ;
118 ;
119 ;
120 ;
121 ;
122 ;
123 ;
124 ;
125 ;
126 ;
127 ;
128 ;
129 ;
130 ;
131 ;
132 ;
133 ;
134 ;
135 ;
136 ;
137 ;
138 ;
139 ;
140 ;
141 ;
142 ;
143 ;
144 ;
145 ;
146 ;
147 ;
148 ;
149 ;
150 ;
151 ;
152 ;
153 ;
154 ;
155 ;
156 ;
157 ;
158 ;
159 ;
160 ;
161 ;
162 ;
163 ;
164 ;
165 ;
166 ;
167 ;
168 ;
169 ;
170 ;
171 ;
172 ;
173 ;
174 ;
175 ;
176 ;
177 ;
178 ;
179 ;
180 ;
181 ;
182 ;
183 ;
184 ;
185 ;
186 ;
187 ;
188 ;
189 ;
190 ;
191 ;
192 ;
193 ;
194 ;
195 ;
196 ;
197 ;
198 ;
199 ;
200 ;
201 ;
202 ;
203 ;
204 ;
205 ;
206 ;
207 ;
208 ;
209 ;
210 ;
211 ;
212 ;
213 ;
214 ;
215 ;
216 ;
217 ;
218 ;
219 ;
220 ;
221 ;
222 ;
223 ;
224 ;
225 ;
226 ;
227 ;
228 ;
229 ;
230 ;
231 ;
232 ;
233 ;
234 ;
235 ;
236 ;
237 ;
238 ;
239 ;
240 ;
241 ;
242 ;
243 ;
244 ;
245 ;
246 ;
247 ;
248 ;
249 ;
250 ;
251 ;
252 ;
253 ;
254 ;
255 ;
256 ;
257 ;
258 ;
259 ;
260 ;
261 ;
262 ;
263 ;
264 ;
265 ;
266 ;
267 ;
268 ;
269 ;
270 ;
271 ;
272 ;
273 ;
274 ;
275 ;
276 ;
277 ;
278 ;
279 ;
280 ;
281 ;
282 ;
283 ;
284 ;
285 ;
286 ;
287 ;
288 ;
289 ;
290 ;
291 ;
292 ;
293 ;
294 ;
295 ;
296 ;
297 ;
298 ;
299 ;
300 ;
301 ;
302 ;
303 ;
304 ;
305 ;
306 ;
307 ;
308 ;
309 ;
310 ;
311 ;
312 ;
313 ;
314 ;
315 ;
316 ;
317 ;
318 ;
319 ;
320 ;
321 ;
322 ;
323 ;
324 ;
325 ;
326 ;
327 ;
328 ;
329 ;
330 ;
331 ;
332 ;
333 ;
334 ;
335 ;
336 ;
337 ;
338 ;
339 ;
340 ;
341 ;
342 ;
343 ;
344 ;
345 ;
346 ;
347 ;
348 ;
349 ;
350 ;
351 ;
352 ;
353 ;
354 ;
355 ;
356 ;
357 ;
358 ;
359 ;
360 ;
361 ;
362 ;
363 ;
364 ;
365 ;
366 ;
367 ;
368 ;
369 ;
370 ;
371 ;
372 ;
373 ;
374 ;
375 ;
376 ;
377 ;
378 ;
379 ;
380 ;
381 ;
382 ;
383 ;
384 ;
385 ;
386 ;
387 ;
388 ;
389 ;
390 ;
391 ;
392 ;
393 ;
394 ;
395 ;
396 ;
397 ;
398 ;
399 ;
400 ;
401 ;
402 ;
403 ;
404 ;
405 ;
406 ;
407 ;
408 ;
409 ;
410 ;
411 ;
412 ;
413 ;
414 ;
415 ;
416 ;
417 ;
418 ;
419 ;
420 ;
421 ;
422 ;
423 ;
424 ;
425 ;
426 ;
427 ;
428 ;
429 ;
430 ;
431 ;
432 ;
433 ;
434 ;
435 ;
436 ;
437 ;
438 ;
439 ;
440 ;
441 ;
442 ;
443 ;
444 ;
445 ;
446 ;
447 ;
448 ;
449 ;
450 ;
451 ;
452 ;
453 ;
454 ;
455 ;
456 ;
457 ;
458 ;
459 ;
460 ;
461 ;
462 ;
463 ;
464 ;
465 ;
466 ;
467 ;
468 ;
469 ;
470 ;
471 ;
472 ;
473 ;
474 ;
475 ;
476 ;
477 ;
478 ;
479 ;
480 ;
481 ;
482 ;
483 ;
484 ;
485 ;
486 ;
487 ;
488 ;
489 ;
490 ;
491 ;
492 ;
493 ;
494 ;
495 ;
496 ;
497 ;
498 ;
499 ;
500 ;
501 ;
502 ;
503 ;
504 ;
505 ;
506 ;
507 ;
508 ;
509 ;
510 ;
511 ;
512 ;
513 ;
514 ;
515 ;
516 ;
517 ;
518 ;
519 ;
520 ;
521 ;
522 ;
523 ;
524 ;
525 ;
526 ;
527 ;
528 ;
529 ;
530 ;
531 ;
532 ;
533 ;
534 ;
535 ;
536 ;
537 ;
538 ;
539 ;
540 ;
541 ;
542 ;
543 ;
544 ;
545 ;
546 ;
547 ;
548 ;
549 ;
550 ;
551 ;
552 ;
553 ;
554 ;
555 ;
556 ;
557 ;
558 ;
559 ;
560 ;
561 ;
562 ;
563 ;
564 ;
565 ;
566 ;
567 ;
568 ;
569 ;
570 ;
571 ;
572 ;
573 ;
574 ;
575 ;
576 ;
577 ;
578 ;
579 ;
580 ;
581 ;
582 ;
583 ;
584 ;
585 ;
586 ;
587 ;
588 ;
589 ;
590 ;
591 ;
592 ;
593 ;
594 ;
595 ;
596 ;
597 ;
598 ;
599 ;
600 ;
601 ;
602 ;
603 ;
604 ;
605 ;
606 ;
607 ;
608 ;
609 ;
610 ;
611 ;
612 ;
613 ;
614 ;
615 ;
616 ;
617 ;
618 ;
619 ;
620 ;
621 ;
622 ;
623 ;
624 ;
625 ;
626 ;
627 ;
628 ;
629 ;
630 ;
631 ;
632 ;
633 ;
634 ;
635 ;
636 ;
637 ;
638 ;
639 ;
640 ;
641 ;
642 ;
643 ;
644 ;
645 ;
646 ;
647 ;
648 ;
649 ;
650 ;
651 ;
652 ;
653 ;
654 ;
655 ;
656 ;
657 ;
658 ;
659 ;
660 ;
661 ;
662 ;
663 ;
664 ;
665 ;
666 ;
667 ;
668 ;
669 ;
670 ;
671 ;
672 ;
673 ;
674 ;
675 ;
676 ;
677 ;
678 ;
679 ;
680 ;
681 ;
682 ;
683 ;
684 ;
685 ;
686 ;
687 ;
688 ;
689 ;
690 ;
691 ;
692 ;
693 ;
694 ;
695 ;
696 ;
697 ;
698 ;
699 ;
700 ;
701 ;
702 ;
703 ;
704 ;
705 ;
706 ;
707 ;
708 ;
709 ;
710 ;
711 ;
712 ;
713 ;
714 ;
715 ;
716 ;
717 ;
718 ;
719 ;
720 ;
721 ;
722 ;
723 ;
724 ;
725 ;
726 ;
727 ;
728 ;
729 ;
730 ;
731 ;
732 ;
733 ;
734 ;
735 ;
736 ;
737 ;
738 ;
739 ;
740 ;
741 ;
742 ;
743 ;
744 ;
745 ;
746 ;
747 ;
748 ;
749 ;
750 ;
751 ;
752 ;
753 ;
754 ;
755 ;
756 ;
757 ;
758 ;
759 ;
760 ;
76
```

Query Match	8.6%	Score 35.2	DB 3	Length 3476
Best Local Similarity	55.4%	Pred. No. 1.7		
Matches 67	Conservative 0	Mismatches 54	Indels 0	Gaps 0
QY 264	caggcgacatgcatactgattgggggaacgaagtgtacgaacaacagagagggag	323		
Db 2143	CTGGCGAAGGTCACCCACCCAGCAGCTGAAGAGGGCGGGAAGAGATCCGGTCA	2202		
QY 324	aagcgcatggaggaacgcggcctnctgtatctcttgcagaggggtctctggaggaagag	383		
Db 2203	GAGAACAGGAAGAGATACATCATCTGCTGACTGACTGGCGTTTCACCCGAGGCGTGGAA	2262		
QY 384	g 384			
Db 2263	G 2263			

```

US-09-010-928B-1
: Sequence 1, Application US/09010928B
: Patent No. 5994099
:
: GENERAL INFORMATION:
: APPLICANT: Lewis, Randolph V
: APPLICANT: Hayashi, Cheryl Y
: TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
: TITLE OF INVENTION: CODING THEREFOR
: NUMBER OF SEQUENCES: 29
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
: STREET: 8110 GATEHOUSE RD. SUITE 500E
: CITY: FALLS CHURCH
: STATE: VIRGINIA
: COUNTRY: UNITED STATES OF AMERICA
: ZIP: 22042
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/010,928B
: FILING DATE: 22-JAN-1998
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Murphy Jr., Gerald M
: REGISTRATION NUMBER: 28977
: REFERENCE/DOCKET NUMBER: 1447-109P
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2830 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: not relevant
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: FEATURE:
:
: NAME/KEY: -
: LOCATION: 1..2830
: OTHER INFORMATION: /note="Flagelliform DNA sequence
: OTHER INFORMATION: taken from the 5' region. The putative start codon is at
: OTHER INFORMATION: position 219"
:
: NAME/KEY: CDS
: LOCATION: 219..2830
:
US-09-010-928B-1

```

Query Match:	8.54;	Score 35;	DB 2;	Length 2830;
Best Local Similarity	44.98;	Pred. No. 1.8;		
Matches 131;	Conservative	0;	Mismatches 161;	Indels 0;
			Gaps	
QY	119	gacgaagaaacattccgcltcaagcttccatgaccccgctcccttgtaagtcacat	178	
Db	900	GGAGGAGGTGAACTTTTCGACACGAGAGTGGCTGGAGGACATACAGCCAGGTGAGCT	959	
QY	179	tcgaagaccacgaggtgtgtatcatgctcgtcctcaagggccgagcgccagttgacacagcat	238	
Db	960	GGACCTGGTGGTGCAGGTGGAGGCTTATGACACGAGAGTGTAGAACTGTGTGAGCCGGA	1015	
QY	239	cggagcgtccgacgtcaacttggtagcagagggcgacgtcatcatcatgatlbggggaaacgaatg	298	
Db	1020	CCAGGAGGTTACGGACCTGTGTGGAGCCGGACACGAGGTTATATGACCTGTGTGACCCGGA	1079	
QY	299	gtacgaagaacacaggaagagaggaagaagacatgagagagagcgccgacchccttatactcttg	358	
Db	1080	CCAGGAGGTTACGGACCTGTGTGGAGCTGTGACCAAGAGGTTACGACACTGTGTGACCTGGG	1138	
QY	359	catgggtgtctctgtgaggaagagaggtgtcccgcaagagcccgcttccttgatg	410	
Db	1140	CTGTGAGGTTACGGACCTGTGTGGAGCTGTGAGGTTACGAGGCTGTGTG	1191	


```
RESULT 12
US-08-676-967-5
; Sequence 5, Application US/08676967
; Patent No. 5747317
; GENERAL INFORMATION:
; APPLICANT: COLLINS, KATHLEEN
; TITLE OF INVENTION: Human Telomerase
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Science & Technology Law Group
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,967
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCB96-055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415)343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2277 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-676-967-5

Query Match      8.4%; Score 34.4; DB 1; Length 2277;
Best Local Similarity 47.5%; Pred. No. 2.4;
Matches 135; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 41 catgcagagaagaaggtcccatctctgtgtacacagagaacccccaagtgcagca 100
   || || || || || || || || || || || || || || || || || || || ||
Db 504 CAAGGCCCTGAAGGCAATGACATGAAGAGATCAAGGGCCGACCGTGGCGTGACTG 563

QY 101 gaccgtgcagcgtgcagagcagaagaaccccttcgcgttcacgttcacacccgc 160
   || || || || || || || || || || || || || || || || || || || ||
Db 564 GGCCGTGGCCAAAGGACAAATGACAGACACCCAGAGCGGTGAGCGCCATCGCGAGAGAA 623

QY 161 gtcccttgtgaagtcacattcgaagcagcggtgtatcatgtctgtcaagccgcgc 220
   || || || || || || || || || || || || || || || || || || || ||
Db 624 GAGCCACG--AGAGCAAGCACACAGAGAGCGGTGAAGAGAGGCCCGCGAGAGAGAA 680

QY 221 gccagttgaccagaccctcgcagcgtcgcagctcacttggagcgagcgtcatcat 280
   || || || || || || || || || || || || || || || || || || || ||
Db 681 CATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 740

QY 281 cgatgtggtggaacagatgtgtacgaagaacacgagagagagagaga 324
   || || || || || || || || || || || || || || || || || || || ||
Db 741 CGACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784

RESULT 13
US-08-676-974-5
; Sequence 5, Application US/08676974
; Patent No. 570422
; GENERAL INFORMATION:
; APPLICANT: COLLINS, KATHLEEN
```

```
; TITLE OF INVENTION: Human Telomerase
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Science & Technology Law Group
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,974
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UCB96-055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415)343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2277 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-676-974-5

Query Match      8.4%; Score 34.4; DB 1; Length 2277;
Best Local Similarity 47.5%; Pred. No. 2.4;
Matches 135; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 41 catgcagagaagaaggtcccatctctgtgtacacagagaacccccaagtgcagca 100
   || || || || || || || || || || || || || || || || || || || ||
Db 504 CAAGGCCCTGAAGGCAATGACATGAAGAGATCAAGGGCCGACCGTGGCGTGACTG 563

QY 101 gaccgtgcagcgtgcagagcagaagaaccccttcgcgttcacgttcacacccgc 160
   || || || || || || || || || || || || || || || || || || || ||
Db 564 GGCCGTGGCCAAAGGACAAATGACAGACACCCAGAGCGGTGAGCGCCATCGCGAGAGAA 623

QY 161 gtcccttgtgaagtcacattcgaagcagcggtgtatcatgtctgtcaagccgcgc 220
   || || || || || || || || || || || || || || || || || || || ||
Db 624 GAGCCACG--AGAGCAAGCACACAGAGAGCGGTGAAGAGAGGCCCGCGAGAGAGAA 680

QY 221 gccagttgaccagaccctcgcagcgtcgcagctcacttggagcgagcgtcatcat 280
   || || || || || || || || || || || || || || || || || || || ||
Db 681 CATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 740

QY 281 cgatgtggtggaacagatgtgtacgaagaacacgagagagagagaga 324
   || || || || || || || || || || || || || || || || || || || ||
Db 741 CGACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784

RESULT 14
US-09-098-487-5
; Sequence 5, Application US/09098487
; Patent No. 5917025
; GENERAL INFORMATION:
; APPLICANT: COLLINS, KATHLEEN
; TITLE OF INVENTION: Human Telomerase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Science & Technology Law Group
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
```



```

? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/804,439A
? FILING DATE: February 21, 1997
? CLASSIFICATION: 424
? ATTORNEY/AGENT INFORMATION:
? NAME: Halle, Lisa A.
? REGISTRATION NUMBER: 38,347
? REFERENCE/DOCKET NUMBER: 09176/004001
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (619) 678-5070
? TELEFAX: (619) 678-5099
? TELEX:
? INFORMATION FOR SEQ ID NO: 13:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2713 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? IS-08-804-439A-13

```

Query Match	8.38;	Score 34;	DB 3;	Length 2713;
Best Local Similarity	47.28;	Pred. No. 3.2;		
Matches 103;	Conservative 0;	Mismatches 115;	Indels 0;	Gaps 0

QY	74	caacagagacacactcccaagctgtagcagagacogtgcacgctgccaagcgccagaggaact	133
Db	1482	CAAGACCACCTCTCCATGAGATTGCGCCGGCTGCAGTTTAACTACAAACACATACACG	1541
QY	134	tcccgcttaagcgttccatgagcccgcgctcttctgtaagtcacatccagaagccaggt	193
Db	1542	CCATGTCAACGATATGTTGGCGCGCTTCCATTCGGGTGTGCAGCTGCAGATCACGA	1601
QY	194	gtgtgatcatgctgtcgaagcgcgcgccagcttgaccagacatcgcgacgtctgcagc	253
Db	1602	GCTGACCTCTGTGAAGAGAGGCCCGCCAGCTGAACCCAAAGCCGATCGCTGGCCACGT	1661
QY	254	tcaactgtgagcaggggagactgcatcatcgatgtgggga	291
Db	1662	GGCGCGCGGAGTGAAGCGCGGAGATCTCGGCACACGTGA	1699

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 07:56:21 ; Search time 796.36 Seconds

(without alignments)
493.713 Million cell updates/sec

Title: US-09-300-482-27

Perfect score: 229

Sequence: 1 cagaccattttctctc.....ctcaattatgggagacaa 229

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802:**

1:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:**
2:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:**
3:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:**
4:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:**
5:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:**
6:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:**
7:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:**
8:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:**
9:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:**
10:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:**
11:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:**
12:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:**
13:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:**
14:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:**
15:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:**
16:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:**
17:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:**
18:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:**
19:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:**
20:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:**
21:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:**
22:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:**
23:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:**
24:	/SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74.2	32.4	1753	21 AAC37505	Arabidopsis thaliana
2	59.2	25.9	369	21 AAC38311	Zea mays DNA fragm
3	58.6	25.6	466	21 AAC41032	Zea mays DNA fragm
4	57.4	25.1	483	21 AAC38393	Zea mays DNA fragm
5	55.8	24.4	477	21 AAC35766	Zea mays DNA fragm
6	50.8	22.2	1639	21 AAC46021	Arabidopsis thaliana
7	47.6	20.8	1639	21 AAC39123	Arabidopsis thaliana
8	33.6	14.7	1105	19 AA296449	S. pneumoniae deri
9	33.2	14.5	2659	24 AB199829	Mouse ischaemic co

10	32.4	14.1	490	22 AAK67424	Human immune/haema
11	32.4	14.1	910715	20 AAX20248	Borrelia burgdorfe
12	32	14.0	385	21 AAC19560	Human secreted pro
13	32	14.0	15914	22 AAK84889	Human immune/haema
14	31.6	13.8	7151	24 AAT70693	Alfalfa acetyl-CoA
15	31	13.5	311	21 AAF08430	Fusarium venenatum
16	30.8	13.4	273254	21 AAC81914	Chlamydia pneumonia
17	30.4	13.3	902	22 AAT194223	Human neuroblastom
18	30.4	13.3	1026	21 AAL51822	Staphylococcus aur
19	30.4	13.3	4924	23 ABL20684	Drosophila melanog
20	30.4	13.3	5116	23 ABL05900	Drosophila melanog
21	30.4	13.3	15249	18 AAT74413	Staphylococcus aur
22	30.2	13.2	450	22 ABA13950	Human nervous syst
23	30.2	13.2	740	22 AAT07506	Human CDNA clone (
24	29.8	13.0	2764	22 AAS42481	Human cDNA encodin
25	29.8	13.0	4434	20 AAK61678	B. burgdorferi ant
26	29.8	13.0	4485	20 AAK61677	B. burgdorferi ant
27	29.8	13.0	116277	20 AAX20249	Borrelia burgdorfe
28	29.6	12.9	564	23 AAT75577	DNA encoding novel
29	29.6	12.9	584	22 ABA63142	Human foetal liver
30	29.6	12.9	584	22 ABA30391	Probe #8857 for ge
31	29.6	12.9	584	22 AAK11574	Human brain expres
32	29.6	12.9	584	22 AAK37343	Human bone marrow
33	29.6	12.9	584	22 AAT18173	Probe #8106 for ge
34	29.6	12.9	584	22 AAT18188	Probe #11874 used
35	29.6	12.9	1435	17 AAT43709	Human inhibitor of
36	29.6	12.9	2286	23 AAS87959	DNA encoding novel
37	29.6	12.9	2580	18 AAT70838	Human apoptosis in
38	29.6	12.9	2589	18 AAT61590	Human c-IAP1. Hom
39	29.6	12.9	3516	22 AAK54169	S. epidermidis gen
40	29.6	12.9	3516	18 AAT72711	Human inhibitor of
41	29.6	12.9	3532	20 AAK22143	Human cellular inh
42	29.6	12.9	3732	19 AAV55040	Human HIRAP-2 codin
43	29.6	12.9	3773	22 AAK54437	S. epidermidis gen
44	29.6	12.9	5608	23 ABL13474	Drosophila melanog
45	29.4	12.8	388	20 AAV90010	EST clone C2326.

ALIGNMENTS

RESULT 1

AC37505 standard; DNA: 1753 BP.

ID	AC37505	standard; DNA: 1753 BP.
XX	AC	AC37505;
XX	DT	17-OCT-2000 (first entry)
XX	DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 17627.
XX	KW	Hybridisation assay; genetic mapping; gene expression control;
XX	KW	Protein identification; signal transduction pathway;
XX	KW	metabolic pathway; promoter; termination sequence; ss.
OS	OS	Arabidopsis thaliana.
XX	PN	EP1033405-A2.
XX	PD	06-SEP-2000.
XX	PF	25-FEB-2000; 2000EP-0301439.
XX	PR	25-FEB-1999; 99US-0121825.
XX	PR	05-MAR-1999; 99US-0123180.
XX	PR	09-MAR-1999; 99US-0123548.
XX	PR	23-MAR-1999; 99US-0125788.
XX	PR	25-MAR-1999; 99US-0126264.
XX	PR	29-MAR-1999; 99US-0126785.
XX	PR	01-APR-1999; 99US-0127462.
XX	PR	06-APR-1999; 99US-0128234.
XX	PR	08-APR-1999; 99US-0128714.
XX	PR	16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132046.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135639.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142290.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143452.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.

}

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147316.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144335
PR	20-JUL-1999	9905-0144332
PR	20-JUL-1999	9905-0144632
PR	20-JUL-1999	9905-0144884
PR	21-JUL-1999	9905-0144881
PR	21-JUL-1999	9905-0145086
PR	21-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145085
PR	22-JUL-1999	9905-0145192
PR	23-JUL-1999	9905-0145145
PR	23-JUL-1999	9905-0145215
PR	23-JUL-1999	9905-0145224
PR	26-JUL-1999	9905-0145276
PR	27-JUL-1999	9905-0145913
PR	27-JUL-1999	9905-0145918
PR	27-JUL-1999	9905-0145919
PR	28-JUL-1999	9905-0145913
PR	02-AUG-1999	9905-0145951
PR	02-AUG-1999	9905-0145386
PR	02-AUG-1999	9905-0145388
PR	03-AUG-1999	9905-0145389
PR	03-AUG-1999	9905-0147038
PR	04-AUG-1999	9905-0147204
PR	04-AUG-1999	9905-0147302
PR	05-AUG-1999	9905-0147192
PR	05-AUG-1999	9905-0147260
PR	06-AUG-1999	9905-0147303
PR	06-AUG-1999	9905-0147416
PR	09-AUG-1999	9905-0147493
PR	09-AUG-1999	9905-0147935
PR	10-AUG-1999	9905-0148171
PR	11-AUG-1999	9905-0148314
PR	12-AUG-1999	9905-0148341
PR	13-AUG-1999	9905-0148365
PR	16-AUG-1999	9905-0148684
PR	17-AUG-1999	9905-0149368
PR	18-AUG-1999	9905-0149175
PR	25-AUG-1999	9905-0149426
PR	27-AUG-1999	9905-0151065
PR	27-AUG-1999	9905-0151066
PR	27-AUG-1999	9905-0151080
PR	30-AUG-1999	9905-0151303
PR	01-SEP-1999	9905-0151348
PR	01-SEP-1999	9905-0151930
PR	10-SEP-1999	9905-0152363
PR	13-SEP-1999	9905-0153070
PR	15-SEP-1999	9905-0153758
PR	16-SEP-1999	9905-0154018
PR	20-SEP-1999	9905-0154039
PR	22-SEP-1999	9905-0154779
PR	23-SEP-1999	9905-0155149
PR	24-SEP-1999	9905-0155486
PR	28-SEP-1999	9905-0155659
PR	29-SEP-1999	9905-0156458
PR	04-OCT-1999	9905-0156596
PR	05-OCT-1999	9905-0157753
PR	06-OCT-1999	9905-0157865
PR	07-OCT-1999	9905-0158029
PR	08-OCT-1999	9905-0158232
PR	12-OCT-1999	9905-0158369
PR	13-OCT-1999	9905-0159293
PR	13-OCT-1999	9905-0159295
PR	13-OCT-1999	9905-0159294

PR	14-OCT-1999;	9905-0159329.
PR	14-OCT-1999;	9905-0159330.
PR	14-OCT-1999;	9905-0159331.
PR	14-OCT-1999;	9905-0159637.
PR	14-OCT-1999;	9905-0159638.
PR	18-OCT-1999;	9905-0159584.
PR	21-OCT-1999;	9905-0160741.
PR	21-OCT-1999;	9905-0160767.
PR	21-OCT-1999;	9905-0160768.
PR	21-OCT-1999;	9905-0160770.
PR	21-OCT-1999;	9905-0160814.
PR	21-OCT-1999;	9905-0160815.
PR	22-OCT-1999;	9905-0160980.
PR	22-OCT-1999;	9905-0160981.
PR	22-OCT-1999;	9905-0160989.
PR	25-OCT-1999;	9905-0161404.
PR	25-OCT-1999;	9905-0161405.
PR	25-OCT-1999;	9905-0161406.
PR	26-OCT-1999;	9905-0161359.
PR	26-OCT-1999;	9905-0161360.
PR	26-OCT-1999;	9905-0161361.
PR	28-OCT-1999;	9905-0161920.
PR	28-OCT-1999;	9905-0161992.
PR	28-OCT-1999;	9905-0161993.
PR	29-OCT-1999;	9905-0162142.

Query Match	25.9%	Score 59.2	DB 21	Length 369
Best Local Similarity	66.4%	Fred. No. 3.2e-09		
Matches 99	Conservative	0	Mismatches 49	Indels 1
			Gaps	1
Qy	65	acaaagaatagcctctgctgcatctgtgctgttaatggcaaatctgccaactaatattgct	124	
Db	177	acaaagaatcgtctctgtccttgctgcttgctgctgcatatnggcgaacctgcccctcaacttgca	236	
Qy	125	tgaagaaggcttcccaattcgggttaacacaggaaccttccaagtatttggccata	184	
Db	237	gagaaaggag-ttcccacatctctgttacaacagcaaacctccaagtgtgacgagacgt	295	
Qy	185	agacgaagcaaacccaggaaggaaccttc	213	
Db	296	gcagcgtgcgaagcagaagaagaaccttc	324	
RESULT	3			
AAC41032				
ID	AAC41032	standard; DNA; 466 BP.		
XX				
AC	AAC41032;			
XX				
DT	17-OCT-2000	(first entry)		
XX				
DE	Zea mays DNA fragment SEQ ID NO: 30405.			
XX				
KM	Hybridisation assay; genetic mapping; gene expression control;			
KW	protein identification; signal transduction pathway; metabolic;			
XX	pathway; promoter; termination sequence; corn; ss.			
XX				
OS	Zea mays subsp. mays.			
XX				
PN	EP1033405-A2.			
XX				
PD	06-SEP-2000.			
XX				
PF	25-FEB-2000; 2000EP-0301439.			
XX				
PR	25-FEB-1999; 99US-0121825.			
PR	05-MAR-1999; 99US-0123180.			
PR	09-MAR-1999; 99US-0123548.			
PR	23-MAR-1999; 99US-0125768.			
PR	25-MAR-1999; 99US-0126264.			
PR	29-MAR-1999; 99US-0126785.			
PR	01-APR-1999; 99US-0127462.			
PR	06-APR-1999; 99US-0128234.			
RR				

PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139489.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140693.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151303.
PR 30-AUG-1999; 99US-0151438.
PR 31-AUG-1999; 99US-0151930.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.

PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	23-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	28-OCT-1999;	99US-0162142.

query match	25.68;	Score 58.6;	DB 21;	Length 466;
-------------	--------	-------------	--------	-------------

```
Matches    99;  Conservatives    0;  Mismatches    49;  Indels    1;  Gaps    1;
```

Oy	65	aacagaataggcctctgcgattgcgcttataatgggaaaatctcggaaccatattgct	124
Db	188	acaagaatcggctctctgcgcctgcgggtcaatgggagaacctctgcccacaattgca	247
Oy	125	tgaaaaggctcccaatccggttaacaagaagaaacattccaaagtatttggcgata	184
Db	248	gagaaaggg-ttcccacatctctgtatacaaaagacaacctcaaggttgagcagacct	306
Oy	185	agacgaagcaaacaccagaagaacacctc	213
Db	307	gcagctgtccaaaggcagaagaacacctc	335

RESULT	4
AAC38393	
ID AAC38393	standard; DNA; 483 BP.

AAC38393;

DT 17-OCT-2000 (first entry)

Zea mays DNA fragment SEQ ID NO: 20807.

KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.

... Zea mays subsp. mays.

PN EP1033405-A2.

06-SEP-2000
PD

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 09-MAR-1999; 99US-0123548.

PR 25-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0127462.

PR	08-APR-1999;	9905-0128234
PR	06-APR-1999;	9905-0128711
PR	16-APR-1999;	9905-0129845
PR	19-APR-1999;	9905-0130047
PR	21-APR-1999;	9905-0130449
PR	23-APR-1999;	9905-0130510
PR	23-APR-1999;	9905-0130891
PR	28-APR-1999;	9905-0131449
PR	30-APR-1999;	9905-0132048
PR	04-MAY-1999;	9905-0132407
PR	05-MAY-1999;	9905-0132484
PR	06-MAY-1999;	9905-0132485
PR	06-MAY-1999;	9905-0132486
PR	07-MAY-1999;	9905-0133487
PR	11-MAY-1999;	9905-0133863
PR	14-MAY-1999;	9905-0134256
PR	14-MAY-1999;	9905-0134218
PR	14-MAY-1999;	9905-0134219
PR	14-MAY-1999;	9905-0134221
PR	14-MAY-1999;	9905-0134370
PR	18-MAY-1999;	9905-0134768
PR	19-MAY-1999;	9905-0134941
PR	20-MAY-1999;	9905-0135121
PR	21-MAY-1999;	9905-0135353
PR	25-MAY-1999;	9905-0136329
PR	25-MAY-1999;	9905-0136629
PR	27-MAY-1999;	9905-0136021
PR	28-MAY-1999;	9905-0136392
PR	01-JUN-1999;	9905-0136782
PR	01-JUN-1999;	9905-0137222
PR	03-JUN-1999;	9905-0137528
PR	04-JUN-1999;	9905-0137502
PR	07-JUN-1999;	9905-0137724
PR	08-JUN-1999;	9905-0138094
PR	10-JUN-1999;	9905-0138054
PR	10-JUN-1999;	9905-0138540
PR	14-JUN-1999;	9905-0138847
PR	16-JUN-1999;	9905-0139311
PR	16-JUN-1999;	9905-0139345
PR	17-JUN-1999;	9905-0139453
PR	18-JUN-1999;	9905-0139454
PR	18-JUN-1999;	9905-0139455
PR	18-JUN-1999;	9905-0139456
PR	18-JUN-1999;	9905-0139345
PR	18-JUN-1999;	9905-0139345
PR	18-JUN-1999;	9905-0139345
PR	21-JUN-1999;	9905-0139817
PR	22-JUN-1999;	9905-0140389
PR	23-JUN-1999;	9905-0140353
PR	23-JUN-1999;	9905-0140354
PR	28-JUN-1999;	9905-0140695
PR	29-JUN-1999;	9905-0140823
PR	30-JUN-1999;	9905-0140921
PR	01-JUL-1999;	9905-0141287
PR	01-JUL-1999;	9905-0141842
PR	01-JUL-1999;	9905-0142154
PR	02-JUL-1999;	9905-0142055
PR	06-JUL-1999;	9905-0142390
PR	08-JUL-1999;	9905-0142803
PR	12-JUL-1999;	9905-0142920
PR	12-JUL-1999;	9905-0142977
PR	13-JUL-1999;	9905-0143542
PR	14-JUL-1999;	9905-0143624
PR	15-JUL-1999;	9905-0144005
PR	16-JUL-1999;	9905-0144085
PR	19-JUL-1999;	9905-0144086
PR	19-JUL-1999;	9905-0144331

	PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
	PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
	PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
	PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
	PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
	PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
	PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
	PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
	PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
	PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
	PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
	PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160769.
	PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160770.
	PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160814.
	PR	23-JUL-1999;	99US-0145145.	PR	21-OCT-1999;	99US-0160815.
	PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160980.
	PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160981.
	PR	26-JUL-1999;	99US-0145276.	PR	22-OCT-1999;	99US-0160989.
	PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161404.
	PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161405.
	PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
	PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161359.
	PR	28-JUL-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161360.
	PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
	PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
	PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161982.
	PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
	PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.
	PR	04-AUG-1999;	99US-0147302.			
	PR	05-AUG-1999;	99US-0147192.			
	PR	05-AUG-1999;	99US-0147260.			
	PR	06-AUG-1999;	99US-0147303.			
	PR	06-AUG-1999;	99US-0147416.			
	PR	09-AUG-1999;	99US-0147493.			
	PR	09-AUG-1999;	99US-0147935.			
	PR	10-AUG-1999;	99US-0148171.			
	PR	11-AUG-1999;	99US-0148319.			
	PR	12-AUG-1999;	99US-0148341.			
	PR	13-AUG-1999;	99US-0148565.			
	PR	13-AUG-1999;	99US-0148684.			
	PR	16-AUG-1999;	99US-0149368.			
	PR	17-AUG-1999;	99US-0149175.			
	PR	18-AUG-1999;	99US-0149426.			
	PR	20-AUG-1999;	99US-0149722.			
	PR	20-AUG-1999;	99US-0149723.			
	PR	20-AUG-1999;	99US-0149929.			
	PR	23-AUG-1999;	99US-0149902.			
	PR	23-AUG-1999;	99US-0149930.			
	PR	25-AUG-1999;	99US-0150566.			
	PR	26-AUG-1999;	99US-0150884.			
	PR	27-AUG-1999;	99US-0151065.			
	PR	27-AUG-1999;	99US-0151066.			
	PR	27-AUG-1999;	99US-0151080.			
	PR	30-AUG-1999;	99US-0151303.			
	PR	31-AUG-1999;	99US-0151438.			
	PR	01-SEP-1999;	99US-0151930.			
	PR	07-SEP-1999;	99US-0152363.			
	PR	10-SEP-1999;	99US-0153070.			
	PR	13-SEP-1999;	99US-0153758.			
	PR	15-SEP-1999;	99US-0154018.			
	PR	16-SEP-1999;	99US-0154039.			

PR 01-APR-1999; 99US-012746Z.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0124284.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0142526.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0142211.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138084.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-014187.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0135139.
PR 23-SEP-1999; 99US-0135145.
PR 24-SEP-1999; 99US-0135486.
PR 28-SEP-1999; 99US-015659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.


```
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-015586.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
```

```
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 22.28; Score 50.8; DB 21; Length 1639;
Best Local Similarity 62.38; Pred. No. 2.7e-06;
Matches 96; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

```
QY 64 aacaagaataggccttctgattggtttaatggcaaatctggcactcaatttc 123
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 77 atctcgatagtgctcgcggtctgcagtcacgggacaaacctgccttaacatcgc 136
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 124 tgaagaaggcttcacatccggttaacaacggaacatttccaagtattggccat 183
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 137 -cgataaagattcccatctccgtctacatcgaaaccttccaagtgcgaacact 195
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 184 aagaagaaacaaacccggaagaaaccttcaatt 217
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 196 tagatcgtgccctccacaggaanaactccagct 229
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 7
AAC39123
ID AAC39123 standard; DNA; 1639 BP.
XX
AC AAC39123;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23456.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000BP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149920.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151087.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0152070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157665.
PR 07-OCT-1999; 99US-0158029.

PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159350.
PR 14-OCT-1999; 9905-0159351.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161922.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

Query Match 20.8%; Score 47.6; DB 21; Length 1639;
Best Local Similarity 61.0%; Pred. No. 2.8e-05;
Matches 94; Conservative 0; Mismatches 59; Indels 1; Gaps 1;

QY 64 aacaagatggccttgctggtatgttgaatgggcaaatctggcaactcaatattgc 123
DB 79 atccgcgactcgctcgcgcgctcgcagatcatggaacaaactcgcttaaacatcgc 138
QY 124 ttgaaaaggcttcccaatccggttaacaacggaacatttccaagtattatggccat 183
DB 139 -cgataaaggattcccaatctcgcttaacaatcgaaacactcccaagtcgacgaaact 197
QY 184 aagacgaagcaaacacggaaggaacattcaatt 217
DB 198 tagatcgctccccaacggaaggaacactcccaat 231

RESULT 8
AAZ96449/C
ID AAZ96449 standard; DNA; 1105 BP.
XX
AC AAZ96449;
XX

DT 10-APR-2000 (first entry)

DE S. pneumoniae derived DNA from ORF #277.

XX Treatment; prevention; disease; diagnosis; gene therapy; screening;
KW bacterial; antimicrobial; antibiotic; pathogenesis; infection; ss.
XX

OS Streptococcus pneumoniae.

XX
PN WO9806734-A1.

XX
PD 19-FEB-1998.

XX
PF 15-AUG-1997; 97WO-US14436.

XX
PR 16-AUG-1996; 96US-0024022.

XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX

PI Black MT, Hodgson JE, Knowles DJC, Lometto MA, Nicholas RO;
PI Stodola RK;
XX
DR WPI, 1998-159452/14.
XX P-PSDB; AAY86134.
PT Streptococcus pneumoniae proteins and related DNA - useful for
screening compounds for antibacterial activity
XX
PS Claim 4; Page 291; 640bp; English.
XX
CC This invention describes novel isolated Streptococcus pneumoniae
CC polynucleotides (see AAZ96173-296494) and their encoded proteins (see
CC AAY85792-Y86182). The DNA, vectors and host cells described in the
CC method of the invention are useful for the recombinant expression of the
CC polypeptides. The polypeptides are useful for treatment or prevention of
CC disease, or diagnosis of disease related to expression or activity of
CC such a polypeptide. They can also be used to screen for compounds which
CC interact with and inhibit or activate such a polypeptide. The
CC polypeptides (or DNA encoding them, via gene therapy) are also useful
CC for inducing an immunological response in a mammal. The antagonists are
CC particularly useful to identify antimicrobial compounds and antibiotics.
CC They are also useful to determine their role in pathogenesis of
CC infection, dysfunction and disease.
XX

SQ Sequence 1105 BP; 303 A; 244 C; 193 G; 365 T; 0 other;

Query Match 14.7%; Score 33.6; DB 19; Length 1105;
Best Local Similarity 52.1%; Pred. No. 0.77;
Matches 75; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 65 acaagaatagccttgctggtatgttgaatgggcaaatctggcaactcaatattgc 124
DB 284 ACTGCTCTGGCTTGGCTTGAAGCGGATTTTCTGCTAAACACCTCCACAAACCT 225
QY 125 ttgaaaaggcttcccaatccggttaacaacggaacatttccaagtattatggccata 184
DB 224 GTCAAGAGATTCGATTTCGAGGCTGAAGTCGACGCTTTCAGAGATTGTGATCCGA 165
QY 185 agacgaagcaaacacggaaggaana 208
DB 164 AAGGAAGTGACAGAGAGAGAGAAA 141

RESULT 9
ABI99829
ID ABI99829 standard; cDNA; 2659 BP.
XX
AC ABI99829;
XX

DT 07-MAR-2002 (first entry)

DE Mouse ischemic condition related cDNA sequence SEQ ID NO:948.

XX Mouse; ischemia; compressive ischemia; occlusive ischemia;
KW vasospastic ischemia; ischemic condition; ischemic disease; ss.
XX

OS Mus musculus.

XX
PN WO200188188-A2.

XX
PD 22-NOV-2001.

XX
PF 18-MAY-2001; 2001WO-JP04192.

XX
PR 18-MAY-2000; 2000JP-0145977.

XX
PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX

DR MPI: 2002-034733/04.
P-PSDB; ABB57340.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
genes -
XX
XX Claim 2: Page 2399+2403; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (1) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (1). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
SQ Sequence 2659 BP; 634 A; 773 C; 544 G; 708 T; 0 other;

Query Match 14.5%; Score 33.2; DB 24; Length 2659;
Best Local Similarity 53.0%; Pred. No. 1.4;
Matches 71; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 5 cctatttttctgtcattgtcctcaattcagagagatgaatgctgcaaccaca 64
DB 1922 cccaacttcattatgtacatcttaaaatgacacacatgagagaaagctaaacaac 1981
QY 65 acaagatagcgcttgctgctgattgctgttaaggcaaatctgacaccatattgct 124
DB 1982 aaaaaaaacatgtaagggtgagcagtaattgcttcattccacatacatctgt 2041
QY 125 tgaagaagcgctcc 138
DB 2042 gtaaacgatttcc 2055

RESULT 10
AAK67424
ID AAK67424 standard; DNA; 490 BP.
XX
AC AAK67424;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SRQ ID NO:22236.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216980.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226581.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229545.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234977.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0234984.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249287.
PR 17-NOV-2000; 2000US-0249289.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX PT useful for preventing, diagnosing and/or treating cancers and
XX PT metastasis -
XX
XX
PS Disclosure; SEQ ID NO 22236; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and

CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 490 BP; 164 A; 75 C; 103 G; 148 T; 0 other;
Query Match 14.1%; Score 32.4; DB 22; Length 490;
Best Local Similarity 64.9%; Pred.No.1.4;
Matches 48; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 40 agattaattatgcgtcaacccacaagaatagagccttgctggtatgctgttaatg 99
Db 402 aaaaaattatgagcatgcatactacaaatgctagctggaaggggtccatg 461
QY 100 gcaaatctggcgc 113
Db 462 gctaaatctgggac 475
RESULT 11
AAK20248/C
ID AAK20248 standard; DNA; 910715 BP.
XX
AC AAK20248;
XX
DT 04-MAY-1999 (first entry)
XX
XX Borrelia burgdorferi polynucleotide sequence #1.
DE
XX
XX Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
XX infection; diagnosis; characterisation; detection; ds.
OS Borrelia burgdorferi.
XX
XX
XX W09858943-A1.
XX
XX
XX 30-DEC-1998.
XX
XX 18-JUN-1998; 98WO-US12764.
XX
XX 03-SEP-1997; 97US-0057483.
XX 20-JUN-1997; 97US-0050359.
XX 22-JUL-1997; 97US-0053344.
XX 22-JUL-1997; 97US-0053377.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX (MDT-) MEDIMUNE INC.
XX
XX Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
XX
XX White OR;
XX
XX WPI; 1999-081217/07.
XX
XX
XX New isolated Borrelia burgdorferi nucleic acids - used to develop
XX products for the detection, diagnosis, characterisation, prevention
XX and therapy of infections, particularly Lyme disease
XX
XX
PS Claim 1; Page 157-671; 1128pp; English.
XX
XX AAK20248 to AAK20402 represent polynucleotide sequences isolated from
CC Borrelia burgdorferi (Bb). Products derived from Bb can be used for

Oy	54	ctcaaccacaacagaatagcttgccttgatgtgcctgtaacagggaacattgcac	113
Db	5807	ctgattctgcagaccagaagcgcccaagcagatattggaatttcaacagaagaactccac	5866
Oy	114	tcaatattgcttgaagaggcttcccaattccggttcaacacgggaacattccaag	171
Db	5867	tttcaattatgcgaactcgtgagagcttltcaagttgcgaacaaggaccttltgaag	5924
RESULT 15			
ID	AAF08430		
AC	AAF08430 standard; CDNA; 311 BP.		
DT	13-MAR-2001 (first entry)		
DE	Fusarium venenatum EST SEQ ID NO:953.		
XX	Multiple gene expression; filamentous fungal cell; EST;		
KW	expressed sequence tag; Fusarium venenatum; Aspergillus niger;		
KW	Aspergillus oryzae; Trichoderma reesei; identification; recombination;		
KW	culture condition; environmental stress; spore morphogenesis;		
KW	metabolic pathway engineering; catabolic pathway engineering; ss.		
XX			
OS	Fusarium venenatum.		
PN	WO200056762-A2.		
PD	28-SEP-2000.		
PF	22-MAR-2000; 2000WO-US07781.		
XX			
PR	22-MAR-1999; 99US-0273623.		
PA	(NOVO) NOVO NORDISK BIOTECH INC.		
PA	(NOVO) NOVO NORDISK AS.		
PI	Berka RM, Rey MW, Shuster JR, Kaupinen S, Clausen IG, Olsen PB;		
DR	WPI; 2000-594572/56.		
XX			
XX	Monitoring differential expression of genes in filamentous fungal cells		
PT	uses fluorescence-labeled nucleic acids isolated from the cells and a		
PT	substrate of expressed sequence tags -		
PS	Claim 86; Page 746; 3161pp; English.		
CC	The present invention describes a method for monitoring differential		
CC	expression of genes in a first filamentous fungal (FF) cell relative to		
CC	expression of the same genes in one or more second filamentous fungal		
CC	cells. The method uses fluorescence-labeled nucleic acids isolated from		
CC	the FF cells and a substrate of expressed sequence tags (EST). The ESTs		
CC	are used in the methods for monitoring differential expression of genes		
CC	in a first filamentous fungal (FF) cell relative to expression of the		
CC	same genes in one or more second filamentous fungal cells. Monitoring		
CC	the global expression of genes from FF cells allows the production		
CC	potential of the microorganisms to be improved. New genes may be		
CC	discovered, possible functions of unknown open reading frames can be		
CC	identified and gene copy number variation and stability can be		
CC	monitored. The expression of genes can be used to study how FF cells		
CC	adapt to changes in culture conditions, environmental stress, spore		
CC	morphogenesis, recombination, metabolic or catabolic pathway		
CC	engineering. Using ESTs provides several advantages over genomic or		
CC	random cDNA clones including elimination of redundancy as one spot on an		
CC	array equals one gene or open reading frame, and organisation of the		
CC	microarrays based on function of the gene products to facilitate		
CC	analysis of the results. AAF07478 to AAF11247 represents ESTs from		
CC	Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus		
CC	niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and		
CC	AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are		
CC	all specifically claimed in the present invention.		

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 07:41:08 ; Search time 185.98 seconds
(without alignments)
302.452 Million cell updates/sec

Title: US-09-300-482-27

Perfect score: 229

Sequence: 1 cagacattattttctgtc.....cttcattatggggaacaa 229

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCUS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30.4	13.3	1026	US-09-149-624-1	Sequence 1, Appl
2	29.6	12.9	1435	PCT-US95-05922A-1	Sequence 1, Appl
3	29.6	12.9	2580	US-08-511-485-7	Sequence 7, Appl
4	29.6	12.9	2589	US-08-569-749-1	Sequence 1, Appl
5	29.6	12.9	2589	PCT-US96-12860-1	Sequence 1, Appl
6	29.6	12.9	3532	US-09-205-204-1	Sequence 1, Appl
7	29.6	12.9	3732	US-09-212-971-7	Sequence 7, Appl
8	29.6	12.9	3732	US-08-800-929A-7	Sequence 7, Appl
9	29.6	12.9	3732	US-09-617-053A-7	Sequence 7, Appl
10	29.6	12.9	72604	US-09-268-992-7	Sequence 7, Appl
11	28.8	12.6	618	US-09-385-982-59	Sequence 59, Appl
12	28.4	12.4	270	US-08-589-080-3	Sequence 3, Appl
13	28.4	12.4	1629	US-08-589-080-2	Sequence 2, Appl
14	27.4	12.0	964	US-08-755-587-23	Sequence 23, Appl
15	27.4	12.0	1041	US-08-755-587-22	Sequence 22, Appl
16	27.4	12.0	1050	US-08-755-587-21	Sequence 21, Appl
17	27.4	12.0	1521	US-08-913-462-37	Sequence 37, Appl
18	27.4	12.0	1599	US-08-737-336-5	Sequence 5, Appl
19	27.4	12.0	6203	US-09-134-218-3	Sequence 3, Appl
20	27.4	12.0	8010	US-09-521-526-2	Sequence 2, Appl
21	27.4	12.0	8010	PCT-US95-11859-2	Sequence 2, Appl
22	27.4	11.8	756	US-08-530-165-1	Sequence 1, Appl
23	27.4	11.8	756	US-07-904-071-1	Sequence 1, Appl
24	27.4	11.8	1008	US-07-904-071-1	Sequence 1, Appl
25	27.4	11.8	1770	US-08-381-603-1	Sequence 1, Appl
26	27.4	11.8	1770	US-08-924-376-1	Sequence 1, Appl
27	27.4	11.8	1770	US-08-685-212-1	Sequence 1, Appl

28	27	11.8	1770	5	PCT-US94-02414-1	Sequence 1, Appl
29	27	11.8	1770	5	PCT-US96-08899-1	Sequence 1, Appl
30	27	11.8	3011	4	US-07-821-716-1	Sequence 1, Appl
31	26.8	11.7	801	4	US-09-276-531-111	Sequence 11, App
32	26.6	11.6	714	3	US-08-643-704A-44	Sequence 44, Appl
33	26.6	11.6	773	4	US-09-227-357-47	Sequence 47, Appl
34	26.6	11.6	816	3	US-08-643-704A-47	Sequence 47, Appl
35	26.6	11.6	900	3	US-08-643-704A-46	Sequence 46, Appl
36	26.6	11.6	1236	3	US-08-643-704A-48	Sequence 48, Appl
37	26.6	11.6	2854	4	US-08-936-165A-66	Sequence 66, Appl
38	26.6	11.6	4649	6	5183745-1	Patent No. 5183745
39	26.6	11.6	5118	4	US-08-669-785-3	Sequence 3, Appl
40	26.6	11.6	6441	4	US-08-669-785-1	Sequence 1, Appl
41	26.6	11.6	6443	6	5183745-5	Patent No. 5183745
42	26.6	11.6	9432	1	US-08-277-231A-1	Sequence 1, Appl
43	26.6	11.6	9432	2	US-08-473-750-4	Sequence 4, Appl
44	26.6	11.6	9432	2	US-08-477-326-4	Sequence 4, Appl
45	26.6	11.6	17041	1	US-08-076-011-1	Sequence 1, Appl

ALIGNMENTS

```

RESULT 1
US-09-149-624-1
; Sequence 1, Application US/09149624
; Patent No. 6222026
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; APPLICANT: Kosmatka, Anna L.
; APPLICANT: Traini, Christopher M.
; APPLICANT: Warren, Richard L.
; TITLE OF INVENTION: gcp
; FILE REFERENCE: GM10166
; CURRENT APPLICATION NUMBER: US/09/149,624
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1023)
US-09-149-624-1

Query Match      13.3%; Score 30.4; DB 4; Length 1026;
Best Local Similarity 52.3%; Pred. No. 0.46;
Matches 67; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 32 attcaggaataatgatgcgtcaaccacacagaataagccttgatgct 91
    || || || || || || || || || || || || || || || || ||
Db 886 attccagatcctaataatgatgcagataatgctgaatgagcgttcgcgcacat 945

QY 92 gttaacggcaaaatctgacacataatgcttgaagaagcctcccaatccggtta 151
    || || || || || || || || || || || || || || || || ||
Db 946 ttgtatcagaagtcgattgctgattgattgacataatgycacagcatatagattta 1005

QY 152 caacggaa 159
    || || || || || || || || || || || || || || || || ||
Db 1006 gaagagta 1013

RESULT 2
PCT-US95-05922A-1
; Sequence 1, Application PC/TUS9505922A
; GENERAL INFORMATION:
; APPLICANT: HE, ET AL.
; TITLE OF INVENTION: Human Inhibitor of Apoptosis Gene 1
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:

```



```

? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/511,485
? FILING DATE: 04-AUG-1995
? CLASSIFICATION: 514
? ATTORNEY/AGENT INFORMATION:
? NAME: Clark, Paul T.
? REGISTRATION NUMBER: 30,162
? REFERENCE/DOCKET NUMBER: 07540/002001
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617/542-5070
? TELEFAX: 617/542-8906
? TELEX: 200154
? INFORMATION FOR SEQ ID NO: 7:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2580 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
? TOPOLOGY: both
? MOLECULE TYPE: DNA (genomic)
? OS-08-511-485-7

```

Query Match	12.9%	Score 29.6	DB 2	Length 2580
Best Local Similarity	47.8%	Pred. No. 1.3		
Matches	86	Conservative	0	Mismatches 94
				Indels 0
				Gaps 0
QY	8	tattttttctgcatctgtcttcaaatccagagataataatcgcgtcaaccacaca	67	
Db	2006	TATATGCCAGGAATGTGCCCTCTCTACAGAAATGCCCATTTTGGAGGGTAAATCA	2065	
QY	68	agaaatgacctgctggatgctgttaatggtggcaaaatctgacatcaaatgtcttga	127	
Db	2006	AGGGTACGTGTGTGCATTTCTCTTAAAGAAAATAGTCTATATTATTAACTGCATTA	2125	
QY	128	aagggcttcccaatccggttaaacagagacatttccaagttctatggtggccaataga	187	
Db	2126	AAAGGCTTTAAATATTGTGTTGAACACTTTGAAGCCATCTAAAGTAAAGGGAATTATGA	2185	

RESULT 4
US-08-569-749-1
Sequence 1, Application US/08569749
Patent No. 6187557
GENERAL INFORMATION:
APPLICANT: Rothe, Mike
APPLICANT: Goeddel, David V
TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FLEHR, HOEBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/569,749
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Brezner, David J.
REGISTRATION NUMBER: 24,774
REFERENCE/DOCKET NUMBER: A-62464/DJB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)781-1989

TELEFAX: (415)398-3249
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2589 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-569-749-1

Query Match 12.9%; Score 29.6; DB 4; Length 2589;
Best Local Similarity 47.8%; Pred. No. 1.3;
Matches 86; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 8 tatttttctgcatctgctcaattcaggagataattatgctcaaccacaaca 67
DB 1963 TACTATGCCAGGAATGTGCCCTTCTCTAAGAAATGCCCTATTTCAGGGATATATCA 2022
QY 68 agaatagccttgctgattgctgttaatgggcaaatctggcactcaatatgtctga 127
DB 2023 AGGTAAGTCTGTGTAATTTCTCTTAAAGAAATAGCTATATTAAACCTGCATFA 2082
QY 128 aaaggctcccaattccggttaacacggaaccattccaaagttatgggcataaga 187
DB 2083 AAAGCTTTAAATTTGTTGAACACTTGAAAGCCATCTAAAGTAAAGGGAATTATGA 2142

RESULT 5
PCT-US96-12860-1

; Sequence 1, Application PC/TUS9612860
; GENERAL INFORMATION:
; APPLICANT: TULARIK, INC.
; TITLE OF INVENTION: INHIBITORS OF APOPTOSIS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/12860
; FILING DATE: 06 AUG 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Serial Nos. 08/512,946 & 08/569,749
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brezner, David J.
; REGISTRATION NUMBER: 24,774
; REFERENCE/DOCKET NUMBER: A-62464/DUB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)781-1989
; TELEFAX: (415)398-3249
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2589 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US96-12860-1

Query Match 12.9%; Score 29.6; DB 5; Length 2589;
Best Local Similarity 47.8%; Pred. No. 1.3;

Matches 86; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 8 tatttttctgcatctgctcaattcaggagataattatgctcaaccacaaca 67
DB 1963 TACTATGCCAGGAATGTGCCCTTCTCTAAGAAATGCCCTATTTCAGGGATATATCA 2022
QY 68 agaatagccttgctgattgctgttaatgggcaaatctggcactcaatatgtctga 127
DB 2023 AGGTAAGTCTGTGTAATTTCTCTTAAAGAAATAGCTATATTAAACCTGCATFA 2082
QY 128 aaaggctcccaattccggttaacacggaaccattccaaagttatgggcataaga 187
DB 2083 AAAGCTTTAAATTTGTTGAACACTTGAAAGCCATCTAAAGTAAAGGGAATTATGA 2142

RESULT 6

US-09-205-204-1
; Sequence 1, Application US/09205204
; Patent No. 5958772
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Elizabeth J. Ackermann
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF CELLULAR INHIBITOR OF APOPTOSIS-1 EXPR
; FILE REFERENCE: RTS-0020
; CURRENT APPLICATION NUMBER: US/09/205,204
; CURRENT FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 3532
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1160)..(3016)
US-09-205-204-1

Query Match 12.9%; Score 29.6; DB 2; Length 3532;
Best Local Similarity 47.8%; Pred. No. 1.5;
Matches 86; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 8 tatttttctgcatctgctcaattcaggagataattatgctcaaccacaaca 67
DB 2928 tagtaagcaggaatgtgccccctctctaagaaatgcccatttgcaggggtataatca 2987
QY 68 agaatagccttgctgattgctgttaatgggcaaatctggcactcaatatgtctga 127
DB 2988 agggtaagctgtgtaattctcttaagaaatagctatatttaacctgcataa 3047
QY 128 aaaggctcccaattccggttaacacggaaccattccaaagttatgggcataaga 187
DB 3048 aaaggctttaaataattgttgaacccttgaagcactcatctaagtaaaagggattatga 3107

RESULT 7

US-09-212-971-7
; Sequence 7, Application US/09212971B
; Patent No. 6107041
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K
; APPLICANT: Pratl, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF IAPs AND
; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; FILE REFERENCE: 07891/009002
; CURRENT APPLICATION NUMBER: US/09/212,971B
; CURRENT FILING DATE: 1998-12-16
; EARLIER APPLICATION NUMBER: 60/017,354


```

? EARLIER FILING DATE: 1996-04-26
? EARLIER APPLICATION NUMBER: 60/030,590
? EARLIER FILING DATE: 1996-11-14
? EARLIER APPLICATION NUMBER: 08/800,929
? EARLIER FILING DATE: 1997-02-13
? NUMBER OF SEQ ID NOS: 17
? SOFTWARE: FastSeq for Windows Version 4.0.
? SEQ ID NO 7
? LENGTH: 3732
? TYPE: DNA
? ORGANISM: Homo sapiens
? OS-09-212-971-7

```

Query Match	Similarity	Score	DB 3	Length	3732
Best Local	47.8%	Pred. No. 1.6			
Matches	86	Conservative	0	Mismatches	94
				Indels	0
				Gaps	0
Oy	8	tatttttcgtcattgctccaatttcaggagatttaattggtctcaaccaca	67		
Db	3148	tagtatgcaggaatgctgcctctctctagaatgccctattgcaagggtata	3207		
Oy	68	agaatagccttgcgcgattgctgttaatgagcaaatctgcaatcatatgctga	127		
Db	3208	agggaactgcttgcatacttctcttaagaanaatagctcatatttcaactgata	3267		
Oy	128	aaaggctcccaattccggttaacaagcaaccattcccaaggtattatggccata	187		
Db	3268	aaagctcttaaatatgcttgtaaacacttgaaacatctaagttaaaagggaattatga	3327		

RESULT 8
 US-08-800-929A-7
 : Sequence 7, Application US/08800929A
 : Patent No. 6133437
 :
 : GENERAL INFORMATION:
 : APPLICANT: Korneluk, Robert G
 : APPLICANT: Mackenzie, Alexander E
 : APPLICANT: Liston, Peter
 : APPLICANT: Baird, Stephen
 : APPLICANT: Tsang, Benjamin K
 : APPLICANT: Pratt, Christine
 : TITLE OF INVENTION: DETECTION AND MODULATION OF
 : TITLE OF INVENTION: IAPS AND NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
 : TITLE OF INVENTION: DISEASE
 : NUMBER OF SEQUENCES: 17
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Clark & Elbing LLP
 : STREET: 176 Federal Street
 : CITY: Boston
 : STATE: MA
 : COUNTRY: USA
 : ZIP: 02110
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FASTSEO for Windows Version 2.0
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/800,929A
 : FILING DATE: 13-FEB-1997
 : CLASSIFICATION: 424
 :
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 60/030,590
 : FILING DATE: 14-NOV-1996
 : APPLICATION NUMBER: 60/017,354
 : FILING DATE: 26-APR-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Bieker-Brady, Kristina
 : REGISTRATION NUMBER:
 : REFERENCE/DOCKET NUMBER: 07891/009001
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 617-428-0200

```

? TELEFAX: 617-428-7045
?
? TELLEX:
?
? INFORMATION FOR SEQ ID NO: 7:
?
?     SEQUENCE CHARACTERISTICS:
?         LENGTH: 3732 base pairs
?         TYPE: nucleic acid
?         STRANDEDNESS: single
?         TOPOLOGY: linear
?
?     MOLECULE TYPE: CDNA
?
? US-08-800-929A-7

```

Query March	12.9%	Score 29.6;	DB 3;	Length 3732;
Best Local Similarity	47.8%;	Pred. No. 1.6;		
Matches	86;	Conservative	0;	Mismatches 94; Indels 0; Gaps 0;
QY	8	tatttttcgctatttgcctcaaatcagaagagataataatgcgtcaaccacaac	67	
Db	3148	TAGATGCCAGSAAATGGCCCTTCTCTNAGAAATGCCCATTTTGCAGGGGATTAATCA	3207	
QY	68	agatagggccttcgtgattgcgtcttaatgycgcaaaatctgycactcaatatigtta	127	
Db	3208	AGGGTACTGTTCTTACATTTCCTCTTAAGAAAAAATAGCTATATATTTAACTGCATTA	3267	
QY	128	aaaggagctcccaatccggttaaagaagacacatctccaagtcttatgggcacaaga	187	
Db	3268	AAAGGCTTTAAATATTTTTGAACACTTGAAGCCATCTAAGTAAAGAAAGGAAATTAATA	3327	

Query Match	12.9%;	Score 29.6;	DB 4;	Length 3732;
Best Local Similarity	47.8%;	Pred. NO. 1.6;		
Matches 86;	Conservative	0;	Mismatches 94;	Indels 0;
Gaps	0			

Query	8	tatttttcgtcatttgccttcaaatctcaggagatttaattgctgcgtcaaccacaaga	67
Dc	3148	tagtatgccaggaatctgcccctctcctaaganaatgccctatttcgagggtataatca	3207
OY	68	agaataggcccttctgtaattgctgtaattgagcaaaatctcgaccacaattatgttga	127
Dc	3208	agggtactcgtctgacattctctcctcttaagaaaaatgctatatttctaactgcataa	3267
OY	128	aaagggtccccaattccggttaacaacaggaacacatttccaaggtattctggccataa	187
Dc	3268	aaagggtcccttaaatctgttgtaaacacttgaagccatctaaagttaaaaaagggatattga	3327


```
RESULT 10
US-09-268-992-7/c
; Sequence 7, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; EARLIER APPLICATION NUMBER: 09/236,134
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n-a, c, g, or t
US-09-268-992-7

Query Match          12.7%; Score 29; DB 4; Length 72604;
Best Local Similarity 61.0%; Pred. No. 9.7;
Matches 47; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 32 attcagagataatgctgctcaaccacacagaatagccttgatgct 91
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 62956 ATTTCAGAGACTTCTTGTCTCCCTCAGCAGAAAAAGAGATAGTAGCAGACACTAGGCT 62897

QY 92 gtaatggcgcaaatct 108
|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 62896 CTTCAGTCTCAATCT 62880

RESULT 11
US-09-385-982-59
; Sequence 59, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: 11
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 59
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(618)
```

```
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-59

Query Match          12.6%; Score 28.8; DB 4; Length 618;
Best Local Similarity 47.1%; Pred. No. 1.3;
Matches 48; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 13 ttctgcatgtgtcattcaattcaggagatgaattatgctcaccacacagaat 72
|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 498 tatcaatcatgtgtcattcaacnaaaaaaaaaaaaaaaaaaaaaaaaaaagga 557

QY 73 aggcctgtcgtgattgctgttaatggcacaacatcgtgcaat 114
|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 558 atctcgtcccgcggtcgttcaaggcnaattcaaccact 599

RESULT 12
US-08-589-080-3
; Sequence 3, Application US/08589080
; Patent No. 5808026
; GENERAL INFORMATION:
; APPLICANT: COHEN, Amos
; APPLICANT: TSE, William
; APPLICANT: ZHU, Weimin
; TITLE OF INVENTION: SEQUENCE OF AFLQ CDNA
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/589,080
; FILING DATE: 23-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9501386.8
; FILING DATE: 23-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rea, Teresa Stanek
; REGISTRATION NUMBER: 30,427
; REFERENCE/DOCKET NUMBER: 024916-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-589-080-3

Query Match          12.4%; Score 28.4; DB 1; Length 270;
Best Local Similarity 62.9%; Pred. No. 1.2;
Matches 44; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 146 gtttaacaacggaaccattccaaagttatgtggccataaagacgaagcaacaggaag 205
||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 117 GGTCAAGACAGACAGCTTGCAAAATGATCGGCAAGCAACATGACAGACACAGAGAA 176

QY 206 aaacttcaa 215
||||  |||
```


TITLE OF INVENTION: Identification and sequencing of the BRCA2 cancer

ADDRESSEE: Bell Seltzer Park & G

ADDRESSEE: Bell Seltzer Park & G

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 05:09:07 ; Search time 5254.53 seconds
(without alignments)
1258.492 Million cell updates/second

Title:	US-09-300-482-225
Perfect score:	316
Sequence:	1 gatagtgcgcacactgag.....gaatgagaggagtacttcga 316

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
```

```

1:  gb_da:*
2:  gb_htg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgq_inv:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

1	160.8	50.9	930	8	AF189365	AF189365	Oryza sat
2	124.4	35.4	131278	2	AC108753	AC108753	Oryza sat
3	100	31.6	922	8	AF370479	AF370479	Arabidops
4	78.4	24.8	89154	8	ATACCT107	ATACCT107	Arabidops
5	69.8	22.1	67712	8	AC022355	AC022355	Arabidops
6	68.4	21.6	5348	8	S77133	S77133	GRI1-genera
7	63	19.9	720	8	CSN03918	CSN03918	Botrytis
8	60.6	19.2	2913	8	ZMU29162	ZMU29162	Zea mays cl
9	60.2	19.1	130843	8	AF464738	AF464738	Zea mays
10	57.8	18.3	424	11	G37533	G37533	SHGC-57943
11	54.8	17.3	1205	6	AX085734	AX085734	Sequence
12	54.8	17.3	1574	9	HSAA4326	HSAA4326	Homo sapi
13	54.8	17.3	1778	9	BC016764	BC016764	Homo sapi
14	54.8	17.3	2130	9	AK056028	AK056028	Homo sapi
15	54.8	17.3	2501	9	BC005148	BC005148	Homo sapi
16	54.4	17.2	178802	2	AL360001	AL360001	Homo sapi
17	53.8	17.0	10304	8	AF165527	AF165527	Zea mays
18	52.2	16.5	3846	8	AB031012	AB031012	Zea mays
19	51.4	16.3	921	3	AY061110	AY061110	Drosophill
20	51.4	16.3	26292	2	AC019942	AC019942	Drosophill
21	51.4	16.3	27521	2	U28991	U28991	Caenorhabdi
22	51.4	16.3	50089	3	AC005641	AC005641	Drosophill
23	51.4	16.3	125150	2	AC005450	AC005450	Drosophill
24	51.4	16.3	173285	3	AC008259	AC008259	Drosophill
25	51.4	16.3	184079	3	AC008258	AC008258	Drosophill
26	51.4	16.3	259764	3	AE003840	AE003840	Drosophill
27	50	15.8	1316	10	BC019126	BC019126	Mus muscu
28	50	15.8	2357	10	BC006953	BC006953	Mus muscu
29	48.4	15.6	2128	8	AF415029	AF415029	Zea mays
30	48.4	15.6	2263	8	AF415109	AF415109	Zea mays
31	48.4	15.6	2291	8	AF415063	AF415063	Zea mays
32	48.4	15.6	3231	8	AF415072	AF415072	Zea mays
33	48.4	15.6	3232	8	AF415080	AF415080	Zea mays
34	48.4	15.6	3271	8	AF415075	AF415075	Zea mays
35	49.4	15.6	3271	8	AF415137	AF415137	Zea mays
36	48.4	15.6	3272	8	AF415044	AF415044	Zea mays
37	48.4	15.6	3272	8	AF415104	AF415104	Zea mays
38	48.4	15.6	3272	8	AF415123	AF415123	Zea mays
39	48.4	15.6	3275	8	AF415030	AF415030	Zea mays
40	49.4	15.6	3276	8	AF415120	AF415120	Zea mays
41	49	15.5	2128	8	AF415029	AF415029	Zea mays
42	49	15.5	2263	8	AF415109	AF415109	Zea mays
43	49	15.5	2263	8	AF415063	AF415063	Zea mays
44	49	15.5	2291	8	AF415072	AF415072	Zea mays
45	49	15.5	3231	8	AF415080	AF415080	Zea mays

ALIGNMENTS

RESULT	1
LOCUS	AF189365
DEFINITION	Oryza sativa D-ribulose-5-phosphate 3-epimerase mRNA, complete cds.
ACCESSION	AF189365
VERSION	AF189365.1
KEYWORDS	GI:6007802
SOURCE	.
ORGANISM	Oryza sativa.
REFERENCE	Oryza sativa
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidaeae; Oryzaceae; Oryza.
TITLE	1 (bases 1 to 930)
JOURNAL	Kopriva, S., Koprivova, A. and Suss, K. H.
MEDLINE	Identification, cloning, and properties of cytosolic D-ribulose-5-phosphate 3-epimerase from higher plants
REFERENCE	J. Biol. Chem. 275 (2), 1294-1299 (2000)
AUTHORS	2 (bases 1 to 930)
TITLE	Kopriva, S., Koprivova, A. and Suss, K. H.
JOURNAL	Direct Submission
	Submitted (20-sep-1999) Institute of Forest Biology and Tree Physiology, Am Flughafen 17, Freilburg im Breisgau 79085, Germany


```

FEATURES
  source
    Location/Qualifiers
      1..930
        /organism="Oryza sativa"
        /db_xref="taxon:4530"
      CDS
        35..721
          /EC_number="5.1.3.1"
          /note="cytosolic isoform"
          /codon_start=1
          /product="D-ribulose-5-phosphate 3-epimerase"
          /protein_id="AF01048.1"
          /db_xref="GI:6007803"
          /translation="MAAARAAKTAIPSMLSDFANLAEADRWRLGADLHMDINDGH
            EPNLTIGAVIQSLKRNHTKAYLDCHLMTNPDYVEPLAKAGASGFTTHIEVSRDNW
            QELI0SIKAKMRGVSLRPGTFVEEFPLVEANENVELVMTVEPFGGOKFEM
            MEKVRALKRKYPPLDIEVDGGLGPTIDVAASAGANCIYAGSSIFGAAPGEYISALR
            KSYEGSQNKS"

BASE COUNT      230 a      206 c      256 g      238 t
ORIGIN

Query Match      50.9%; Score 160.8; DB 8; Length 930;
Best Local Similarity 83.2%; Pred. No. 1.2e-42;
Matches 183; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Oy 1 gataagtgccacacacgagagaagaataccctccctcaacatagaggttgatgtrgtc 60
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 518 GAAAGGTGGCTGCCCTGAGAAAGAAAGTACCCATCCCTTGACATCGAGGTTGATGGTGT 577
Oy 61 ctatgctcctcaacatagacgttgcgcacatctgtctgaggccaattgcatcgtcgtcga 120
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 578 CTGGGCTCTTCACACATGATGATGCGTCGATCTGCGGCTAACTGATTCGCGTGA 637
Oy 121 agcctcatttggcgtctggagacccagagagccatcatctgtctgaggagaagcgtc 180
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 638 AGCTCAATATTGGAGCAGCTGAACCGAGAGAGTCTATTCGCGACCTAAGAGAGCGCTT 697
Oy 181 gaaggctctcagacaaacaaactattgtgtttcgtc 220
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 698 GAAGGATCTCAGAAACAAGCTGATCTGTGTTTGTACT 737

RESULT 2
AC108753 131278 bp DNA linear HMG 31-JAN-2002
LOCUS Oryza sativa chromosome 9 clone OSJNBa0010B06, *** SEQUENCING IN
DEFINITION PROGRESS ***, 5 unordered pieces.
ACCESSION AC108753.1 GI:18449959
VERSION AC108753
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
REFERENCE 1 (bases 1 to 131278)
AUTHORS Yun,D.-W., Hahn,J.-H., Yoon,U.-H., Lee,J.-S., Lee,M.-C., Eun,M.Y.
and Kim,H.-I.
Oryza sativa BAC OSJNBa0010B06 genomic sequence
TITLE Oryza sativa BAC OSJNBa0010B06 genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 131278)
AUTHORS Hahn,J.-H. and Kim,H.-I.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2002) Rice Genome Sequencing Project, National
Institute of Agricultural Science and Technology(NIAST), RDA, 249
Seodun-dong, Suwon 441-707, Korea (E-mail:jnhahn@rda.go.kr,
Tel:82-31-290-0309, Fax:82-31-290-0308)
COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

```

```

* be preserved
* 1 50005: contig of 50005 bp in length
* 50006 50105: gap of unknown length
* 50106 52284: contig of 2179 bp in length
* 52285 52384: gap of unknown length
* 52385 84269: contig of 31885 bp in length
* 84270 84369: gap of unknown length
* 84370 95915: contig of 11546 bp in length
* 95916 96015: gap of unknown length
* 96016 131278: contig of 35263 bp in length.
FEATURES
  source
    Location/Qualifiers
      1..131278
        /organism="Oryza sativa"
        /variety="Nipponbare"
        /db_xref="taxon:4530"
        /chromosome="9"
        /clone="OSJNBa0010B06"

BASE COUNT      35990 a      27760 c      28607 g      38517 t      404 others
ORIGIN

Query Match      39.4%; Score 124.4; DB 2; Length 131278;
Best Local Similarity 82.2%; Pred. No. 4.8e-30;
Matches 143; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Oy 47 aggtgatgtgtgtctcaggctccctcaacatagacgtgcccacatcgttgggccaatt 106
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 57301 AGTTGATGTGTGTCTGGGCTCTTCACCATGATGATGTGCTGCTGCGGCTACT 57360
Oy 107 gcatcgtcgtcggaaagctcattatgttgcgtctgcggacccagagacatcatctgtgc 166
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 57361 GCATTTCCTCTGCAACCTCAATATTGGAGAGCTGCAACGAGAGAGCATATATCGGCAC 57420
Oy 167 tgaaggagagcgctcgcagggcctcagacaaacaaactgatttgggtgtcctc 220
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 57421 TAAGGAAGAGCGTTGAGAGATCTCAGAACAAAGCTGATCTGTGTTTGTACT 57474

RESULT 3
AF370479 922 bp mRNA linear PLN 30-APR-2001
LOCUS Arabidopsis thaliana putative D-ribulose-5-phosphate 3-epimerase
DEFINITION (F28J7.18) mRNA, complete cds.
ACCESSION AF370479
VERSION AF370479.1 GI:13877556
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 922)
AUTHORS Lam,B., Southwick,A., Karlin-Neumann,G., Nguyen,M., Miranda,M.,
Lam,B., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamuya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Natusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
TITLE Submitted (17-APR-2001) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
JOURNAL
COMMENT
e-mail for correspondence: arabsequence.stanford.edu

```

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFI cDNAs (RAFI cDNA: "RIKEN Arabidopsis Full-length cDNA"): Seki,M., Natusaka,M., Ishida,J., Satou,M., Kamuya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFI cDNAs: Lam,B., Karlin-Neumann,G., Nguyen,M., Southwick,A., Miranda,M., Palm,C.J.,

REFERENCE 2 (bases 1 to 67712)
AUTHORS Lin, X. and Kaul, S.
TITLE Direct Submission
JOURNAL Submitted (01-FEB-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, xlinet@igf.org
REFERENCE 3 (bases 1 to 67712)
AUTHORS Town, C.D. and Kaul, S.
TITLE Direct Submission
JOURNAL Submitted (19-JAN-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@igf.org
COMMENT On Jan 19, 2001 this sequence version replaced gi:12280840.
Address all correspondence to: atef@igf.org

BAC clone F9N12 is from Arabidopsis thaliana chromosome 1
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including Genscan+ (Chris Burge,
http://CCR-081.mt.edu/GENSCAN.html), GenemarkHMM (Mark Borodovsky,
http://genemark.biology.gatech.edu/Genemark/), Glimmer4 (a variant
of Glimmer3, see Minaela Perlea,
http://www.tigr.org/softlab/glimmerm.htm/glimmerm.html, and
Mperle@tigr.org), searches of the complete sequence against a
peptide database and the plant EST database at TIGR
(http://www.tigr.org/tdb/tgi.shtml). Annotated genes are named to
indicate the level of evidence for their annotation. Genes with
similarity to other proteins are named after the database hits.
Genes without significant peptide similarity but with EST
similarity are named as unknown proteins. Genes without protein
or EST similarity, that are predicted by more than two gene
prediction programs over most of their length are annotated as
hypothetical proteins. Genes encoding tRNAs are predicted by
tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).
Simple repeats are identified by RepeatMasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html).
Location/Qualifiers

FEATURES

SOURCE

1. 67712
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F9N12"
1. 12863
/note="overlap with BAC clone F2K11
[AC008047:88005..100867]."
misc_feature
1
/note="21552 nt beyond this point were not included in the
submitted sequence due to an overlap with another bac
[F2K11]"
join(3453..4024,4432..4526,4569..4721,4795..4877,
4969..5104,5177..5350,5443..>5622)
/gene="F9N12.1"
3453..5622
/gene="F9N12.1"
/note="contains similarity to Flavin-containing
monooxygenase"
join(3469..4024,4432..4526,4569..4721,4795..4877,
4969..5104,5177..5350,5443..5622)
/gene="F9N12.1"
/codon_start=1
/product="unknown protein: 3469-5622"
/protein_id="AAG52151.1"
/db_xref="GI:12324359"
/translation="MAPALSPTRSHHVAVIGAGPAGIYAARELREHSHVVEKKO
VGCTWITYDEVSDPLSDPTTRSYHSSVYRSLRNGTRCTGTRDFFPVRSQSRD
RRRPSHGEVLAIVLAKFAKEFGIEEMVFEIEVAVSAEIGKMRISTEKKKV
RRDEIYAVVNCNGHYEPLRAQLPGISWPKGEMSHSNRYRPEPRDKVLFIDHNV
AVLIGNSSAEDISHDIAVAKEVAVACRSNPADFFIKOTGVNNIMTSMIESVEDS
SVYVNGKTIISVDITMCTGYKHPPLDTNGIIVVDNRGPLYKDDVEPRFAFWLS
FIGIEMOVLPRPMEIOSKMTAGVLSGRILPLSKEDMMLIKETPTSLVNGCIKPRRYT
HRMGTORFYDNWMLASQGCSETEEMREKMLANGVREKAHPETTYRDMEDHHLVSEA
YQDFSLIS"

mRNA

gene

CDS

CDS

mRNA

gene

<6658..>9312
/gene="F9N12.2"
6658..9312
/gene="F9N12.2"
/note="contains similarity to NBS/LRR disease resistance
protein"
6658..9312
/gene="F9N12.2"
/codon_start=1
/product="unknown protein: 6658-9312"
/protein_id="AAG52149.1"
/db_xref="GI:12324357"
/translation="WGISFSIPDPVCNKSQWMDKVSYTHNEKNIALEKTKEL
KAKRDLERLREKAREARGIOLSEFQWLDVAIVEDITITLLRDNRNEIOLICRF
CSKSLTRSYRYKSVFLRLREYERLKGWVGVITQASTSAFEERLOPTTYGQDTPL
DKAGHLMEDGIGIMGMGCVGKTTTLTOLYNNENDKDGCDFGIWVVSQEEFE
VKODETHAQKIGIGDEPMKOKSKOGKITLYNTRKRSVLPFDITWEKVDLAIEGVPD
PRTKGRKLATFTRSQEVCARMGVHPHEVOCLENNVAFDLFOKRVGQTTIGSDGIP
OLARIYAKKCCGLPLALVIGETSKCRTIOEMRAIIVLNSYAEFTGMEKVLPL
QYSLNLKGEQVSSLYCALYPEDAKILKEDLIEMHICEITIDSEGIERAADGYE
IIGCLVARSALMEMDDGDRAVACHVDVAREMALIASELGKQAFIVRAGVREI
PKIKMNVVRRKSLMENKIHILVSGECMETTLTLGKREGSIRSOLKTIISSEFFN
MPKILAVLDSHKSLFELPEETSNVLSKYLLVLTETSHLPKIGIOELKTIHMLEY
TRKLESTIGSISLHMLKYLKPLRSRPLMDINVKRELPLEHEITITITIDPAKQELS
SHRLSHRLLETIGSSVSSLNHRESVSDKIREQIKSCSIEIKMGICITPFLS
LVDVINECEGLRELTFLFAPKIRLSVWNAKDEDITNEBKAGGEGSLPPEL
NLTVDHDLPKLKIYWRPLPELCLEINIRECPNLRKLPDSTSGKQGENGCIIRNKD
SWFEGVWADATKRRFLPSQILSLK"
<11196..>13892
/gene="F9N12.3"
11196..13892
/gene="F9N12.3"
/note="similar to NBS/LRR disease resistance protein
GB:AAC26125 from [Arabidopsis thaliana]"
11196..13892
/gene="F9N12.3"
/codon_start=1
/product="hypothetical protein: 11196-13892"
/protein_id="AAG52150.1"
/db_xref="GI:12324358"
/translation="MGISFSIPDPVCNKSQWMDKVSYTHNEKNIALEKTKEL
KAKRDLERLREKAREARGIOLSEFQWLDVAIVEDITITLLRDNRNEIOLICRF
CSKSLTRSYRYKSVFLRLREYERLKGWVGVITQASTSAFEERLOPTTYGQDTPL
DKAGHLMEDGIGIMGMGCVGKTTTLTOLYNNENDKDGCDFGIWVVSQEEFE
VKODETHAQKIGIGDEPMKOKSKOGKITLYNTRKRSVLPFDITWEKVDLAIEGVPD
PRTKGRKLATFTRSQEVCARMGVHPHEVOCLENNVAFDLFOKRVGQTTIGSDGIP
OLARIYAKKCCGLPLALVIGETSKCRTIOEMRAIIVLNSYAEFTGMEKVLPL
QYSLNLKGEQVSSLYCALYPEDAKILKEDLIEMHICEITIDSEGIERAADGYE
IIGCLVARSALMEMDDGDRAVACHVDVAREMALIASELGKQAFIVRAGVREI
PKIKMNVVRRKSLMENKIHILVSGECMETTLTLGKREGSIRSOLKTIISSEFFN
MPKILAVLDSHKSLFELPEETSNVLSKYLLVLTETSHLPKIGIOELKTIHMLEY
TRKLESTIGSISLHMLKYLKPLRSRPLMDINVKRELPLEHEITITITIDPAKQELS
SHRLSHRLLETIGSSVSSLNHRESVSDKIREQIKSCSIEIKMGICITPFLS
LVDVINECEGLRELTFLFAPKIRLSVWNAKDEDITNEBKAGGEGSLPPEL
NLTVDHDLPKLKIYWRPLPELCLEINIRECPNLRKLPDSTSGKQGENGCIIRNKD
SWFEGVWADATKRRFLPSQILSLK"

[illegible]

REMARK Genbank staff at the National Library of Medicine created this entry (NCBI gi|563164519) from the original journal article. This sequence comes from Fig. 1, Location/Qualifiers

FEATURES

source 1..5348

organism="Zea mays"

db_xref="taxon:4577"

join(1449..2583,2671..2749,2876..2998,4387..4503,4610..4968)

gene 2177..4669

gene="GRF1"

note="general regulatory factor"

join(2177..2583,2671..2749,2876..2998,4387..4503,4610..4669)

gene="GRF1"

note="14-3-3 protein homolog; This sequence comes from fig. 5"

codon_start=1

product="GF14-6"

protein_id="BAB33304.1"

db_xref="gi:998430"

translation="MASAELSREENVYAKLAEOAEKREEMVEKAKVDSSELTVEERNLVAVKKNVIGARRASWRITISLEQEEGNGEDRYTLIDYDKITETELTKCDGILKLETLHLVPSSPAESKVFYLLKKGDYRYRLAEFGAEKKDAEMTMVAKKQADALAPLHPHRIQLALNFSAFYEIILNSPPRACSLAKQADEAISLDTLSESYKDSLTLMQLRLNMLTWTSSISDPPAEIRKRRDSSSEQ"

BASE COUNT 1325 a 1169 c 1203 g 1651 t

ORIGIN

Query Match 21.6%; Score 68.4; DB 8; Length 5348;

Best Local Similarity 87.2%; Pred. No. 1,7e-11;

Matches 75; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 225 agtactccctccgtttttttttatcgcgcgttttaagttcaacatgacacagcggaag 284

|||||

DB 3413 AATACTCCTCCGTTGTTTATTTATTTGTCATTTAGTTAAATAAATACATGACGGCGC 3472

QY 285 actgatattcgagaatcgagagagata 310

|||||

DB 3473 ACAATATTTCGAGAAATGAGAGTAGTA 3498

RESULT 7

LOCUS CENS01918 720 bp mRNA linear PLN 02-SEP-1999

DEFINITION Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation.

ACCESSION AL111059

VERSION AL111059.1 GI:5825679

KEYWORDS cDNA library; nitrogen deprivation.

SOURCE Botryotinia fuckelliana.

ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes; Helotiales; Sclerotiniaceae; Botryotinia.

REFERENCE 1 (bases 1 to 720)

Aulton, F., Lewis, C., Fortini, D., Pradier, J.M. and Brygoo, Y. Direct Submission

TITLE Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr, 78026 Versailles, France

JOURNAL 2 (bases 1 to 720)

REFERENCE Genoscope.

AUTHORS Direct Submission

TITLE Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage : CP 5706 91057 Evry cedex - FRANCE E-mail : seqref@genoscope.cns.fr

JOURNAL - Web : www.genoscope.cns.fr

COMMENT The cDNA library to be analyzed within the framework of this project was created using a Botrytis cinerea strain which was grown under conditions of nitrogen deprivation, which is the normal situation for B. cinerea during its development on its host plant. The library was produced in an oriented direction, in the pBSII vector.

FEATURES

Location/Qualifiers

GIGFQRCNDVKNLAPPKNLNSFVKAKADPMODNEGILLYPAGYPSKIKIKSRKSHS
GPNHAFMYKSETSSROPTHAPEKKTPIASNDHTISPKTILDAATVILNSGKYAS
GGSSIIIDSCGCTNMTREKRMFSSEYKNOOPQATTFPEDGNGVTFRRSDSIATKFG
VLEBGLIYVDDEBRLDTCLAKTINMGWLMHRRLAHVGKMLKHLKGEHLIGLNVH
FEKDEKSOQTEETKGLFLRAONEFLKIKIRSDNGTEFNSQIEGLEEGIKHEF
SSPYPOONGVVERKNTLDMARTLDEBTPDQFMEAVNTACVAINRYLHILK
KTSVELLGGKKNISYFRRVGSCKTILVKRGRSKFAFKTGVFLGDYDPTRAVRF
NKSGLVEVSCDVVDFETNGSOVEQVDLDENGNDAPVALRNMISIDVCKESELFP
SAOQOPSSSTASPTONEDERAQVDEDOANEPEODGIDGCGDANODKDEDEORP
PAPRHQAIOQDHPDVLIGDIHKGVTRSRVAFCEHSEVSSIEPRHVEALQDSD
VWVAMQELNNEFTRNEVHILVPRNQNVTGRVAFKRNKODEGVTRKARLVAGXS
QVEGLDFEETAPVAPARLESIRILAYATYHGFXYOMDVKSAFLNGPIKEEYVQOP
GFEDSEYNHYKLSKALYGLKQAPRAVYECRLDRLNGRVKAGADPTLFTKLEND
LFVCOIYVDDIIFGSTNESTCEFSRIMTQCEPMSELMELGFGFOYKOICEGFIIS
OTKXTODLSKFGKDAKPIKTPMGTNCHLDLDTGKSVYDQKFIHPPSRRLSRAIM
OSTLGAQGHAPPLPLFPADPTLEFTTPEQASASNNPHASSPSPQSRCPPE"

Query Match 19.1%; Score 60.2; DB 8; Length 130843;
Best Local Similarity 79.8%; Pred. No. 1.7e-08;
Matches 71; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 223 aagtaactccctcggtttttttattcgctgagtttagtccaacatgaaatagcgga 282

DB 119955 AACTACTCCACTCTATCTCTTTTATTTGTCGCGTTTACCAAAAAGAACTATCGGG 120014

QY 283 cgactgatattcgagaatgaggaagctac 311

DB 120015 CAACAATATTTTGAGAACATAGTAGTAC 120043

RESULT 10
LOCUS G37533 424 bp DNA linear STS 31-MAR-1998
DEFINITION SHGC-57943 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION G37533
VERSION G37533.1 GI:2997184
KEYWORDS STS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 424)
AUTHORS Myers, R.M.
TITLE Human STS (1997)
JOURNAL Unpublished
COMMENT

Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu
Primer A: CATGACACGACGACATCGGC
Primer B: TTGAGACACGACGATCCATC
STS size: 232

PCR Profile:

Initial incubation: 95 degrees C for 10 minutes

Denaturation: 94 degrees C for 30 seconds

Annealing: 60 degrees C for 30 seconds

Polymerization: 72 degrees C for 23 seconds

PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9700

Protocol:

Template: 25 ng

Primer: each 1 uM

dNTPs: each 200 uM

Amplifying Gold Polymerase: 0.07 units/ul

Buffer:

MgCl2: 2.5 mM

KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

Prepared with primer pairs derived from W80360 -- UniGene.
Location/Qualifiers
1. 424

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="5"
/clone_11b="Human"

STS
primer_bind
132..363
132..151
primer_bind
complement(344..363)

BASE COUNT 114 a 103 c 75 g 125 t 7 others
ORIGIN

Query Match 18.3%; Score 57.8; DB 11; Length 424;
Best Local Similarity 50.0%; Pred. No. 3.8e-08;
Matches 137; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 4 aagtgcgacacatgagaagaagtaacctccctgacatagagtgatggtgcta 63

DB 375 AAGTTCACCTGGTTGAGAGACCCAGTCCATCTTGATATFAGAGGTGATGAGTA 316

QY 64 ggctcccaacatagacagctgagccgcatctgctgggccaatgcatgctgctggaac 123

DB 315 GGTCTGACACCTGCTCCATTAATGTGACGAGCGAGNGCTAACATGTTGTTGCTGCACT 256

QY 124 tctatatggcgctgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 183

DB 255 GCTATTNTNGNGTGAAGACCCAGATCTGTGATCAATCTATTGAAGATGTTGCTCA 196

QY 184 ggctccgaacaacaacagatttgctgctgctgctgctgctgctgctgctgctgct 243

DB 195 GAAGCTGCTCAGAAAGCTTCTTGTGATCGGTGAACCATTAAGNGCCAGTGTCTCTGT 136

QY 244 ttatatcgctcggttttagttcaaacatgaacta 277

DB 135 CATGAATCTCCCTTTACTGGAACAGAAATA 102

RESULT 11

AX085734 1205 bp DNA linear PAT 09-MAR-2001

LOCUS AX085734
DEFINITION Sequence 12 from Patent WO0112790.
ACCESSION AX085734
VERSION AX085734.1 GI:13275684
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1205)
AUTHORS Bandman, O., Lu, D.A., Yue, H., Tran, B., Hillman, J.L., Baughn, M.R.,
Lal, P. and Tang, Y.T.
TITLE Isomerase proteins
JOURNAL Patent: WO 0112790-A 12 22-FEB-2001;
Incyte Genomics, Inc. (US)

FEATURES
source
1. 1205
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 2481256CB1"

BASE COUNT 331 a 224 c 291 g 359 t
ORIGIN

Query Match 17.3%; Score 54.8; DB 6; Length 1205;
Best Local Similarity 50.0%; Pred. No. 4.7e-07;
Matches 137; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY	4	aaggcgccacacgagagaagaatlaacccctccctctgacatagaaggttgatggtgctta	63
Db	314	AAGTTTCACTGCTGTGGAGGACCACATGTCCTTCATTGGAATATAGAGTGGATGGTGAATA	373
QY	64	ggtccctcaacacatagacgctgcgcacatctctgtggccaattgcatactgcctggaagc	123
Db	374	GGTCTGCACTGTGCCAATTAATGTGCACAGGACAGGACGTAACTGATGTGTCTGGCAGT	433
QY	124	tctatatttgcctgcctgcgagcccgagacatcatatctgtcgtcaggaagaagcctgag	183
Db	434	GCTTTTATGAGGAGTGAAGACCCCGATCTGTGATCAATCTATTAAAGAAATGTTGCTCA	493
QY	184	ggtcttcgaacaanaaatctgattctgtgtttctgcgtctglaaglaatccctccgtttttt	243
Db	494	GAACCTGCTCAGAAACGTTCTTGTGATGGTGAACCATTAAGAGACCCAGTGTTCCTGTT	553
QY	244	tttatctgcgcgtttttagtccaacaatgaacta	277
Db	554	CATGAATATCTCCCTTTTACTGGAAAAACAGATA	587

LOCUS	BC016764	1778	bp	mRNA	linear	PRI 09-NOV-2001
DEFINITION	Homo sapiens, clone MGC:22637 IMAGE:4107315, mRNA, complete cds.					
ACCESSION	BC016764					
VERSION	BC016764.1 GI:16876981					
KEYWORDS	MGC.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
AUTHORS	1 (bases 1 to 1778)					
TITLE	Strausberg, R.					
JOURNAL	Direct Submission					
REMARK	Submitted (05-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA					
COMMENT	NIH-MGC Project URL: http://mgc.ncl.nih.gov Contact: MGC help desk Email: cgaps-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: CLONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbio.org contact: amandansystemsbiology.org Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whitting					
FEATURES	<p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov</p> <p>Series: IRAL Plate: 32 Row: C Column: 6</p> <p>This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.</p> <p>Location/Qualifiers</p> <p>1..1778</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="MGC:22637 IMAGE:4107315"</p> <p>/tissue_type="Bone marrow, acute myelogenous leukemia"</p> <p>/clone_lib="NIH MGC_55"</p> <p>/lab_host="DH10B"</p> <p>/note="Vector: pDNR-LIB"</p> <p>35..721</p>					
CDS						

Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Keltman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 7 Row: e Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction.

FEATURES

source
Location/Qualifiers
1..2501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3505010"
/issue_type="Placenta, chorlocarcinoma"
/clone_lib="NIH_MGC_21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
<1..701
/codon_start=3
/product="Unknown (protein for IMAGE:3505010)"
/protein_id="AAH05148.1"
/db_xref="GI:13477345"
/translation="HEGMAAGCGKIGPSILNSDLANLGAECRLMLDSGADYLHLDVMD
GHFVPIITGHPVYESLRKLGODPRFDMHMYSKPEQVKKPAVAGANQYTHLEAT
ENRGALIKDIRENGKRVGLAIKRGTSYERLAPWANDIDALVTVPEPGGQFEMEDM
MRKVMRLKTFPSLDIEVDGVPDTHKCAEAGANMIVSGSAIMRSEDPRSVINLLR
NVCSEAQRSLDR"

CDS

BASE COUNT 785 a 414 c 526 g 776 t
ORIGIN

Query Match

Best Local Similarity 17.3%; Score 54.8; DB 9; Length 2501;
Matches 137; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 4 aaggtgcgcacactgagaagaagtiacccttccttgacatagaggttgatggtgcta 63
DB 489 AAGGTTCACTGTTGAGAGACCAGTTCCATCTTTGGATATAGAGGTGATGGTGGAGTA 548
QY 64 gttcctcaacatagacgtgcgcgcatctgctggggccaatgcatcgctcgctggaagc 123
DB 549 GGTCTGACACTGTCCATAATGTGCAGAGGAGGAGCTAACATGATGTGCTGGCAGT 608
QY 124 tctatatctggcgtcgagaccagagccatcatatctgctgaggaagagcgtcgag 183
DB 609 GCTATTATGAGGAGTGAGACCCAGATCTGTGATCATCTATTAGAAATGTTGCTCA 668
QY 184 ggtctcagaacaaactgatttggtgtctgctgtaagtaacctccctcgcttttt 243
DB 669 GAAGCTGCTCAGAAACGTTCTCTTGATCGGTGAACCATTAAGAGCCCACTGTTCCGTT 728
QY 244 ttattcgtcgctttagttcaaacatgaacta 277
DB 729 CATGAATCTCCCTTTACTGAAACAGGAAATA 762

Search completed: July 3, 2002, 05:09:52
Job time: 48772 sec


```
RESULT 2
; US-09-328-111-388/C
; Sequence 388, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertli, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 388
; LENGTH: 667
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(667)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-388

Query Match      10.9%; Score 34.4; DB 4; Length 667;
Best Local Similarity 63.1%; Pred. No. 0.013;
Matches 53; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Oy 57 tggctagagcttcacacccatagacgtggcgcatcgtcggtgggccaatgctcgtgc 116
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 352 TGGTGCAGGAGTGTGCACACATAGAGAGTGGCTGTGCTGTGATGGAAAGAGCTTT 293
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Oy 117 tggagctctctatatcttggcgctgc 140
      ||||| ||||| ||||| ||||| |||||
Db 292 TGGCAGCGCTTATCATTTGGTTATGC 269

RESULT 3
; US-09-056-226-1/c
; Sequence 1, Application US/09056226B
; Patent No. 6177614
; GENERAL INFORMATION:
; APPLICANT: Colasanti, Joseph J.
; APPLICANT: Sundaresan, Venkatesan
; TITLE OF INVENTION: Control of Floral Induction in Plants
; FILE REFERENCE: CSHL94-04A4
; CURRENT APPLICATION NUMBER: US/09/056,226B
; CURRENT FILING DATE: 1998-04-07
; EARLIER APPLICATION NUMBER: US 09/000,640
; EARLIER FILING DATE: 1997-12-30
; EARLIER APPLICATION NUMBER: US 08/804,104
; EARLIER FILING DATE: 1997-02-20
; EARLIER APPLICATION NUMBER: PCT/US96/03466
; EARLIER FILING DATE: 1996-03-15
; EARLIER APPLICATION NUMBER: US 08/406,186
; EARLIER FILING DATE: 1995-03-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
```

```
LENGTH: 3693
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (241)...(329)
; OTHER INFORMATION: Intron
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (628)...(746)
; OTHER INFORMATION: Intron
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (921)...(2347)
; OTHER INFORMATION: Intron
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(3693)
; OTHER INFORMATION: n = A, T, C or G
US-09-056-226-1

Query Match      10.8%; Score 34; DB 4; Length 3693;
Best Local Similarity 74.1%; Pred. No. 0.046;
Matches 43; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy 235 ccgctttctttatcgtcgcgctttagtcaacatgacactagcgagactgat 292
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1457 CCTTCTTTTATCAGTGGCATTTAGTTAAATCAATCAGGAGCATTAATAT 1400

RESULT 4
; US-08-737-524B-26
; Sequence 26, Application US/08737524B
; Patent No. 5912414
; GENERAL INFORMATION:
; APPLICANT: CARL SAVERIO FALCO
; APPLICANT: DOMINICK ANTHONY GUIDA, JR.
; APPLICANT: MARY ELIZABETH HARNETT LOCKE
; TITLE OF INVENTION: NUCLEIC ACID FRAGMENTS, CHIMERIC
; TITLE OF INVENTION: GENES AND METHODS FOR INCREASING
; TITLE OF INVENTION: THE METHIONINE CONTENT OF THE SEEDS
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95 (7.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/737,524B
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNN M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1059-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 26:
```



```

; SEQUENCE CHARACTERISTICS:
;   LENGTH: 3639 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;
US-08-737-52AB-26

```

Query Match	10.7%;	Score 33.8;	DB 2;	Length 3639;
Best Local Similarity	68.1%;	Pred. No. 0.054;		
Matches 62;	Conservative 0;	Mismatches 27;	Indels 2;	Gaps 1;

QY 211 tgccttcgtgtaagtaactccctcgcgttttttttlatcgtgcggttttaagttcaaca 270
 | | | | | | | | | | | | | | | | | | | | | |
 Db 55 TAATTAGTTGGATTATATTCATCTTTCTTTTATTTGT--CTGTTTTAGTNAAAA 112

Qy	271	tgaactagcggacgactgatatcagaatg	301
Db	113	TGAACTAACCAACGACCAATATTGAGAACG	143

RESULT 5
US-08-390-878-16/c
; Sequence 16, Application US/08390878
; Patent No. 5700683

APPLICANT: Stover, Charles R.
APPLICANT: Mahlars, Gregory G.
TITLE OF INVENTION: VITROUBLE-ATTENUATING GENETIC DELETIONS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Stewart Street Tower, 20th
STREET: Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,878
FILING DATE: 17-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 15371A-17
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/543/9600
TELEFAX: 415/543/5043
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 16885 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
;

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-390-878-16

```

Query Match	10.1%;	Score 32;	DB 1;	Length 16885;
Best Local Similarity	55.4%;	Pred. NO. 0.55;		
Matches 62;	Conservative	0;	Mismatches 50;	Indels 0;
				Gaps 0;

OY 31 ccttccttgacatagagtltgatgtgctcagtcccttcaaccatagactggccgca 90
|| || || || || || || || || || || || || || || || || || || || ||
Db 5391 ccggcgattgcattggcgctgtggctgatttagatgccggcggtmccggcagaggccggc 5332

QY 91 tctcgctgggccaattgcacgtcgtcygaagctctatatattgccccgtcgg 142
 ||| ||| ||| | ||||| | ||| | ||| | |||
Db 5331 TCTCCTGCGGGCATCGGCATCGAGTTGGCCGAGGCCGAGTGCGGTGGGG 5280

RESULT 6
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A

- ; APPLICANT: FLEISCHMAN, Robert D.
- ; APPLICANT: WHITE, Owen R.
- ; APPLICANT: FRASER, Claire M.
- ; APPLICANT: FRASER, Claire M.

```

1  TITLE OF INVENTION:  DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
2
3  TITLE OF INVENTION:  TUBERCULOSIS
4
5  FILE REFERENCE:  24366-20007 00
6
7  CURRENT APPLICATION NUMBER:  US/09/103,840A
8
9  CURRENT FILING DATE:  1998-06-24

```

```

; NUMBER OF SEQ ID NOS: 2
;
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;
; LENGTH: 4403765
;

```

```

: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: FEATURE:
: OTHER INFORMATION: CDC 1551
: OTHER INFORMATION: "n" bases at various positions throughout the sequence
: OTHER INFORMATION: represent a, t, c or g
: JS-09-103-840A-2

```

Query Match	10.1%;	Score 32;	DB 4;	Length 4403765;
Best Local Similarity	55.4%;	Pred. No. 10;		
Matches 62;	Conservative	0;	Mismatches 50;	Indels 0;
			Gaps	0;

OY 31 ccttccttgatcagatgatgtatggtgcaggcccttcaacatagacgtggcgga 90
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4345576 CCGGCGATGGGCAGGGGGGGGTGGGTCTTTAAATCGCGCGGGTTCGGCGAGGGCGCG 4345512

OY 91 tctgctgggccaatlgtcatcgtcgcygaagactataatttggcgcgttcg 142

Db 4345516 TCTTCCTGCGCGATCGGCATCGGAGTTGGCGGAGCGCGAGGTGGGGCTGGGG 43454655

RESULT 7
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A

```

; Patent NO. 6294328
;
; GENERAL INFORMATION:
;
; APPLICANT: FLEISCHMAN, Robert D.
;
; APPLICANT: WHITE, Owen R.
;
; APPLICANT: FRASER, Claire M.

```

: APPLICANT: VENTER, John C.
 : TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 : TITLE OF INVENTION: TUBERCULOSIS
 :

```

; FILE REFERENCE: 24366-2000/.00
;
; CURRENT APPLICATION NUMBER: US/09/103,840A
;
; CURRENT FILING DATE: 1998-06-24
;
; NUMBER OF SEQ ID NOS: 2

```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529

```

```

; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

Query Match	10.1%;	Score 32;	DB 4;	Length 4411529;
Best Local Similarity	55.4%;	Pred. NO. 10;		
Matches	62;	Conservative	0;	Mismatches 50;
			Indels	0;
			Gaps	0;

Oy 31 ccttcacctgacatagaagttatggttggttctaggtccttcaacatagacgttgccgca 90
Db 4353323 CCGCGCATGGCGCATCGGGGGGCTGGGGTGAATGATGGCGCGGTCGCGGAGAGGGCGC 4353264
Oy 91 tctgtgtgggcgaattgcatcgctgcgtgaaagcttatatttggcgcgtcgca 142
Db 4353263 TCTCTCTGGGGGATCGCGCATCGGATGGAGTTGGCGGAGCGCGAGAGTGGGGGGGGG 4353212

```

1      RESULT      8
2      US-09-276-531-8
3      : Sequence 8, Application US/09276531
4      : Patent No. 6183968
5      : GENERAL INFORMATION:
6      : APPLICANT: Bandman, Olga
7      : APPLICANT: Lal, Preeti
8      : APPLICANT: Hillman, Jennifer L.
9      : APPLICANT: Yue, Henry
10     : APPLICANT: Reddy, Roopa
11     : APPLICANT: Guegler, Karl J.
12     : APPLICANT: Baughn, Mariah R.
13     : TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
14     : NUMBER OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
15     : NUMBER OF SEQUENCES: 134
16     : CORRESPONDENCE ADDRESS:
17     : ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
18     : STREET: 3174 PORTER DRIVE
19     : CITY: PALO ALTO
20     : STATE: CALIFORNIA
21     : COUNTRY: USA
22     : ZIP: 94304
23     :
24     : COMPUTER READABLE FORM:
25     : MEDIUM TYPE: Floppy disk
26     : COMPUTER: IBM PC compatible
27     : OPERATING SYSTEM: PC-DOS/MS-DOS
28     : SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
29     : CURRENT APPLICATION DATA:
30     : APPLICATION NUMBER: US/09/276,531
31     : FILING DATE: Herewith
32     : CLASSIFICATION:
33     : PRIOR APPLICATION DATA:
34     : APPLICATION NUMBER: 60/079,677
35     : FILING DATE: March 27, 1998
36     : CLASSIFICATION:
37     : ATTORNEY/AGENT INFORMATION:
38     : NAME: Lynn E. Murry, Ph.D.
39     : REGISTRATION NUMBER: 42,918
40     : REFERENCE/DOCKET NUMBER: PA-0008 US
41     : TELECOMMUNICATION INFORMATION:
42     : TELEPHONE: (650) 855-0555
43     : TELEFAX: (650) 845-4166
44     : INFORMATION FOR SEQ ID NO: 8:
45     : SEQUENCE CHARACTERISTICS:
46     : LENGTH: 3556 base pairs
47     : TYPE: nucleic acid
48     : STRANDEDNESS: single
49     : TOPOLOGY: linear
50     : IMMEDIATE SOURCE:
51     : LIBRARY: L1YR1UT01
52     : CLONE: 1753826
53     :
54     : US-09-276-531-8

```

	Query Match	9.4%	Score 29.8	DB 4	Length 3556
Best Local Similarity	59.0%	Pred. No. 1.4			
Matches	69	Conservative	0	Mismatches 47	Indels 1
					Gaps 1
QY	4	aagtcgcacactgagaagaagtagccttccttcgaatgaggt-tgtagtgcgtct	62		
Db	531	AAGTCTACTGCTGTGAGACCCAGTCCCAATTGGMATATPAGTCCGATGGAGACT	590		
QY	63	aggtctcaaccatagacgtgtgcgcgcatctgtgtggccaattgcatctgtcgttg	119		

```

DB      591 AGGTCCTGACACTGTCCATTAATATGTCCAGAGGCAGAGCTAACATGATTGTCTGTGG 647

RESULT      9
US-09-488-856A-10
/ Sequence 10, Application US/09488856A
/ Patent No. 6316259
/ GENERAL INFORMATION:
/ APPLICANT: Brett P. Monia
/ APPLICANT: Robert McKay
/ APPLICANT: Madeline M. Butler
/ APPLICANT: Jacqueline Wyatt
/ TITLE OF INVENTION: ANTISENSE MODULATION OF GLYCOGEN SYNTHASE KINASE 3 ALPHA
/ FILE REFERENCE: R5-0115
/ CURRENT APPLICATION NUMBER: US/09/488,856A
/ CURRENT FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 88
/ SEQ ID NO 10
/ LENGTH: 12394
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (115)...(397)
/ NAME/KEY: CDS
/ LOCATION: (2438)...(2625)
/ NAME/KEY: CDS
/ LOCATION: (5639)...(5722)
/ NAME/KEY: CDS
/ LOCATION: (5864)...(5974)
/ NAME/KEY: CDS
/ LOCATION: (7902)...(8032)
/ NAME/KEY: CDS
/ LOCATION: (8121)...(8227)
/ NAME/KEY: CDS
/ LOCATION: (9197)...(9294)
/ NAME/KEY: CDS
/ LOCATION: (9375)...(9470)
/ NAME/KEY: CDS
/ LOCATION: (9898)...(10084)
/ NAME/KEY: CDS
/ LOCATION: (10431)...(10523)
/ NAME/KEY: CDS
/ LOCATION: (11713)...(11786)
/ US-09-488-856A-10

```

```

Query Match          9.48;   Score 29.8;   DB 4;   Length 12394;
Best Local Similarity 66.28;
Matches 43;   Conservative 0;   Mismatches 22;   Indels 0;   Gaps

Oy  155  tcatatctgtcgtgaggaagagcgtcgaaggccctcagaacaacatgtattgtgtc 214
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  1178  taataatgttctaagtgatgtgtctatctgcccctgaataaaaaaagtatgtagggtta 1239
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Oy  215  tctgc 219
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  1238  tctcc 1242

RESULT  10
US-08-581-148C-30
Sequence 30, Application US/08581148C
Patent No. 6060644

GENERAL INFORMATION:
APPLICANT: Schnable, Patrick S.
APPLICANT: Robertson, Donald S.
APPLICANT: Hansen, Joel D.
APPLICANT: Nikolau, Basil J.
APPLICANT: Xu, Xiaojie
APPLICANT: Xia, Yijie

TITLE OF INVENTION: ISOLATION AND USE OF CUTICULAR LIPID
PARENT OF INVENTION: GENES

```


NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leydig, Volt & Mayer, Ltd.
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: IL
COUNTRY: US
ZIP: 60601-6780
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/581,148C
FILING DATE: 29-DEC-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Larcher, Carol
REGISTRATION NUMBER: 35243
REFERENCE/DOCKET NUMBER: 71380
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 616-5600
TELEFAX: (312) 616-5700
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 6343 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Zea mays
INDIVIDUAL ISOLATE: Z. mays Glossy2 locus DNA
US-08-581-148C-30

Query Match
Best Local Similarity 9.2%; Score 29.2; DB 3; Length 6343;
Matches 37; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 224 aagttactccctcggtttttttatctgctgcttttagtcaaatga 273
Db 3878 ACGTACTCCCTCGCTCTTTTATGCTGTTGTCACAGATCA 3927

RESULT 11
US-08-030-096-3
Sequence 3, Application US/08030096
Patent No. 5426041
GENERAL INFORMATION:
APPLICANT: Fajlanski, Steven F.
TITLE OF INVENTION: BINARY CRYPTOCYTOTOXIC METHOD OF HYBRID
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/030,096
FILING DATE: 22-MAR-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/556,917

FILING DATE: 20-JUL-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA91/00255
FILING DATE: 22-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 33229/164/PIHI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 8585 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(308..370, 1136..1261, 6369..6428, 7198..7353)
US-08-030-096-3

Query Match
Best Local Similarity 9.1%; Score 28.8; DB 1; Length 8585;
Matches 63; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 154 atcatatcgtgctgaggaagcgcgtcagagcctcagaacaaacatttgctg 213
Db 6383 ACCATATTTTTCACAAATATAAGTGACCAACGATATCTCATTTTTTTT 7042

QY 214 ctgcgttaagacccctcgtttttttatctgctgcttttagtcaaatga 273
Db 7043 ATATACGTGGCAAAATATTTCTTTTCTTTTACTTATACGTTTAAATGAATGT 7102

RESULT 12
US-07-960-389-1/c
Sequence 1, Application US/07960389
Patent No. 5705611
GENERAL INFORMATION:
APPLICANT: HAYASHIDA, Kasuhiko;
TITLE OF INVENTION: Human GM-CSF Receptor Component
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corporation
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disc
COMPUTER: Apple Macintosh
OPERATING SYSTEM: System Software 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/960,389
FILING DATE: 07-JAN-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 554,745
FILING DATE: 18-JUL-1990
APPLICATION NUMBER: PCT/US 91/04846
FILING DATE: 16-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Blasdale, John H. C.
REGISTRATION NUMBER: 31,895
REFERENCE/DOCKET NUMBER: DX01430
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 298-2902

TELEFAX: (908) 298-5388
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3475 base pairs
TYPE: nucleotides
STRANDEDNESS: single
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: human
FEATURE:
OTHER INFORMATION: DNA sequence encoding Human GM-CSF receptor
US-07-960-389-1

Query Match 8.9%; Score 28.2; DB 1; Length 3475;
Best Local Similarity 49.7%; Pred. No. 5.3;
Matches 72; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 68 ctcaaccatagagctgagccgcatctgctgggccaattgcatcgctcgtaagctcta 127
DB 1254 CTCACATCCTTGTCACGTTGAGGAGTGGAGGGCCATCTGATGTTCACTGAGCTTTTA 1195
QY 128 ttttggcgctcgagaccgaacatactgctgctgaggaagcgctcgagggct 187
DB 1194 TGTGTTTCTCTGCCCTTGGCTGACAGACGATGATTGGCCGTGGGTGGGGGT 1135
QY 188 ctcaagaacaaactgatttggtg 212
DB 1134 CGGCGACGGGAATCTGCGAGTGTGTG 1110

RESULT 13
US-08-714-918-76/c
Sequence 76, Application US/08714918
Patent No. 6037123

GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ying
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,918
FILING DATE: September 13, 1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 222/005
TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 3305 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-714-918-76

Query Match 8.8%; Score 27.8; DB 3; Length 3305;
Best Local Similarity 51.2%; Pred. No. 7.1;
Matches 65; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 180 cgaagctctcaagaacaaactgatttggtgttctcgtcaaaagtactccctcggt 239
DB 789 CTAGGTATTAAGCAACATGCGACACGCTTGATTATAGTCTTGAAGTATGACTTTTAT 730
QY 240 tttttatctcgctggttttagttcaacatgaactagcgagcgactgatatcgagaa 299
DB 729 CATTTGTTTCAAGCGCTAATCTTTGAATAATTACTAGCTGACATATTAATTGCMAAAT 670
QY 300 tgaaggg 306
DB 669 TAAACGG 663

RESULT 14
US-09-265-315-76/c
Sequence 76, Application US/09265315
Patent No. 6187541

GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ying J.
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/265,315
FILING DATE: March 9, 1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/714,918
FILING DATE: September 13, 1996
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 240/247

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 05:09:52 ; Search time 5254.53 Seconds
(without alignments)
1545.237 Million cell updates/sec

Title: US-09-300-482-298

Perfect score: 388
Sequence: 1 ggaagaagaagaagaatg.....ttaagcctgggaccccgct 388

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_vl:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description

1	239.8	61.8	930	8	AF189365	AF189365	Oryza sat
2	226.6	58.5	922	8	AF370479	AF370479	Arabidops
3	122.6	31.6	12091	1	AE006493	AE006493	Streptoco
4	112.8	29.1	10029	1	AE006423	AE006423	Lactococc
5	96.8	24.9	1316	10	BC019126	BC019126	Mus muscu
6	96.8	24.9	2357	10	BC006953	BC006953	Mus muscu
7	95.2	24.5	338579	1	AP003004	AP003004	Mesorhizo
8	91.4	23.6	313450	1	AL596170	AL596170	Listeria
9	90.2	23.2	347050	1	AL591981	AL591981	Listeria
10	88.8	22.9	1205	6	AX085734	AX085734	Sequence
11	88.8	22.9	1778	9	BC016764	BC016764	Homo sapi
12	88.8	22.9	2501	9	BC005148	BC005148	Homo sapi
13	88.8	22.9	3556	6	AR129150	AR129150	Sequence
14	88	22.7	10029	1	AE009551	AE009551	Brucella
15	87.2	22.5	303249	1	AP001515	AP001515	Bacillus
16	85.2	22.0	927	33	AC057311	AC057311	Giardia
17	85.2	22.0	14229	1	AE001286	AE001286	Chlamydia
18	80.8	20.8	11208	1	AE008544	AE008544	Streptoco
19	80.8	20.8	178802	2	AL360001	AL360001	Homo sapi
20	79.8	20.6	303750	1	AP003133	AP003133	Staphyloc
21	79.8	20.6	347235	1	AP003361	AP003361	Staphyloc
22	79.2	20.4	137952	2	SPNEU1909	SPNEU1909	Streptoco
23	79.2	20.4	163443	2	AC006280	AC006280	Plasmodiu
24	79.2	20.4	196149	2	AC004709	AC004709	Plasmodiu
25	79	20.4	911	8	SCPOS18	SCPOS18	S.cerevisia
26	79	20.4	1328	8	SCYJL121C	SCYJL121C	S.cerevisia
27	78.6	20.3	5960	1	RCU23145	RCU23145	Rhodobacter
28	78.2	20.2	921	3	AY061110	AY061110	Drosophil
29	77.6	20.0	12566	6	BD003829	BD003829	Polynucle
30	77.6	20.0	12939	1	AE002306	AE002306	Chlamydia
31	77.6	20.0	14906	1	AE007487	AE007487	Streptoco
32	77.4	19.9	131278	2	AC108753	AC108753	Oryza sat
33	76.4	19.7	10150	1	AE002217	AE002217	Chlamydia
34	76.4	19.7	12973	1	AE001604	AE001604	Chlamydia
35	76.4	19.7	299650	1	AP002545	AP002545	Chlamydia
36	76	19.6	306250	1	SMES91788	SMES91788	Sinorhizo
37	75.4	19.4	849	12	AF070943	AF070943	Expressio
38	75.4	19.4	993	12	AF070942	AF070942	Expressio
39	75.4	19.4	1198	8	SPR5P3E	SPR5P3E	Splinaclia
40	75.4	19.4	1230	8	AF070941	AF070941	Splinaclia
41	75	19.3	67712	8	AC022355	AC022355	Arabidops
42	74.2	19.1	3400	1	TP097573	TP097573	Trepionema
43	74.2	19.1	10081	1	AE001263	AE001263	Trepionema
44	74.2	19.1	39729	8	SPAC31G5	SPAC31G5	S.pombe chr
45	74.2	19.1	296750	1	AP003191	AP003191	Clostridi

ALIGNMENTS

RESULT 1
AF189365 LOCUS 930 bp mRNA linear PLN 10-JAN-2000
AF189365 Oryza sativa D-ribulose-5-phosphate 3-epimerase mRNA, complete cds.
AF189365
AF189365.1 GI:6007802
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa.
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 930)
Kopriva,S., Koprivova,A. and Suss,K.H.
Identification, cloning, and properties of cytosolic
D-ribulose-5-phosphate 3-epimerase from higher plants
J. Biol. Chem. 275 (2), 1294-1299 (2000)
Submitted (20-SEP-1999) Institute of Forest Biology and Tree
Physiology, Am Flughafen 17, Freiburg im Breisgau 79085, Germany

FEATURES	Source	Location/Qualifiers
CDS		1..930 /organism="Oryza sativa" /db_xref="taxon:4530" 35..721 /EC_number="5.1.3.1" /note="cytosolic isoform" /codon_start=1 /product="D-ribulose-5-phosphate 3-epimerase" /protein_id="AA01048.1" /db_xref="GI:6007803" /translation="MAAAAKIAIPSMISDPANLAAPADPRLGADMLMDMDGHEVPNLTITAPVIGISLRKHTKAYLDCHLMTWTPSPDVEPLAKAGSGFTFHEVSRDNNQELDIOSIAKMRPEVSLRPGTPVEEFPVLEAEENPELVMTVPEFGQKMPKMEKRALRKRYPSLDIEVDGLGPSTIDVAASAGANCIVASISIFGAEPGEVTSALRKSVESQSKMS"
BASE COUNT	230 a 206 c 256 g 238 t	
ORIGIN		
Query Match	61.8%;	Score 239.8; DB 8; Length 930;
Best Local Similarity	78.8%;	Pred. No. 3.6e-56;
Matches 286; Conservative	0; Mismatches 77; Indels 0; Gaps 0;	
OY	26	gacaccgaaataagctccttgatgctctcttcgcgacttcgcaatttggctccgagc 85
Db	49	GGCGGCGAAGATGAGCGCGTGCATGCTCTCGTGGACTTCGCCAACCTCGCGCGAGGC 108
OY	86	tcagcgatgctcccacttcgcgcgcgatggctggtccacatgacatcatgattgacatt 145
Db	109	CGACCGCATGTCGCCCTCGCGCGACTGGCTCCACATGACATCATGAGCGGCGACTT 168
OY	146	tgtccccaatttaactatgtgcgcctccagtaattgaaggltgagaaaacacacaagc 205
Db	169	TGTTCTATCTTACTATTATGGAGCTCCAGTATTGAGAGCTTGAGAGACACACCAAGC 228
OY	206	aatattgattgtcaaccttaagtgtacaatcccttgattatgltgaaccttggcaaa 265
Db	229	ATATTGGACGTGCCATCTTATGTGTACCAATCTTCGGATTATGTAGAACCATTTAGCAAA 288
OY	266	agctgtgcttccttggttttaccatctcacgttagagacacacaaaagtaacttgaagaact 325
Db	289	AGCTGTGCTCCTCAGGTTTCACATTCATATAGAACTATCCAGAGACAAATTGGCAAGAACT 348
OY	326	tatccaaagaatcaagtcacatbgtgcatgtattcctgtgtaagcataaagcctbgtgaccc 385
Db	349	CATCCAAAGTATCAAGCAAGGAGGATGCGACCGGGGTATCATATTGAGCCACGCGACTCC 408
OY	386	cgt 388
Db	409	TGT 411
RESULT 2		
FEF70479		
LOCUS		
DEFINITION	AF370479	922 bp mRNA linear PLN 30-APR-2001
ACCESSION	AF370479	Arabidopsis thaliana putative D-ribulose-5-phosphate 3-epimerase (F28J7.18) mRNA, complete cds.
VERSION	AF370479	
KEYWORDS	AF370479.1	GI:13877556
SOURCE	PL1.CDNA.	
ORGANISM	thale cress.	
REFERENCE	Arabidopsis thaliana	
AUTHORS	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
	1 (bases 1 to 922)	
	Lam, B., Southwick, A., Karlin-Neumann, G., Nguyen, M., Miranda, M., Palm, C. J., Bowser, L., Jones, T., Banb, J., Carlinici, P., Chen, H., Cheuk, R., Chung, M. K., Hayashizaki, Y., Ishida, J., Kamiya, A., Kawai, J., Kim, C., Lin, J., Liu, S. X., Narusaka, M., Pham, P. K., Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K., Shinozaki, K., Ecker, J. T., Theologis, A., and Davis, R. W.	

TITLE	Direct Submission					
JOURNAL	Submitted (17-APR-2001) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA					
COMMENT	e-mail for correspondence: arab@sequence.stanford.edu RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : "RIKEN Arabidopsis Full-Length cDNA") : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamuya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K. The Saik, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs. Lam,B., Karlin-Neumann,G., Nguyen,M., Southwick,A., Miranda,M., Palm,C.J., Bowser,L., Jones,F., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologos,A. and Davis,R.W. Lam,B., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs. Location/Qualifiers 1..922 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="III" /clone="RAFLL1-02-F16" /note="This clone is in pBluescript ecotype: Columbia" 1..922 /gene="F28J7.18" 51..728 /gene="F28J7.18" /codon_start=1 /product="putative D-ribulose-5-phosphate 3-epimerase" /protein_id="AAK43856.1" /db_xref="gi:13877557" translation="MVSPIAKIPMSLSSPFANLAEANRMIDLGANMLIMDIDHPFYV NLTIAGPIVESLRKHTNAYLDCHLMVTNPMDYVQOMKAGASGTFHYEVNDMMOOL VERIKSTGMRPGVALLPEGTPEVOYYPVEIGTPEVMILYMVFEGSGOKRPMPOMK VRALROKYPTIIDIQVDGLGPSTIDTAAAGANCIIVAGSVFGAPEDDYISLTRTSV EKADPTT"					
BASE COUNT	230 a	212 c	237 g	243 t		
ORIGIN						
Query Match	58.5% Score 226.8; DB:8 Length 922;					
Best Local Similarity	75.4%; Pred. No. 1.5e-52;					
Matches 282;	Conservative 0;	Mismatches 92;	Indels 0;	Gaps 0;		
Oy	15	aagatggaatgcacgcgaaataagctccttcgaagctctctccgactcgccaattg	74			
Dd	45	AAGAAGATGTTGTCCGCCAATAATAGCACCGTCGATCGACTGGCAATCTA	104			
Oy	75	gtcttcggagtcaagcgcatgctccacttggcgccgatitggtccacaatygaatcatg	134			
Dd	105	GC GGCGGAGGCCAATCGGATGATCATTTGGCGCCCAACTGGCTTCACATGATATTATG	164			
Oy	135	gatggcatttgtccccaatctaacttatggcgctcgaattatgaaagttagaaga	194			
Dd	165	GACGGGCAITTTTGTCCTTAATCTTCAGTTGTGTCTCTGTTCATCGAAGATTTGGCAGAG	224			
Oy	195	cacacaaaggcataattggaattgtacccttaatgyltiacaatatccctctgataitgtgaa	254			
Dd	225	CACACTATATGATATCTTGATTTGGCACCTTATMGTAGAACACCACCATGATGATTAGCTTCA	284			
Oy	255	cccttggaagaagaagcgtgaccttggtttccatttcscgaagaagacataaagaataac	314			
Dd	285	CAGATGGCTTAAGCTGGGGCTTCTGGTTTCAATTTCCACGTCGAGGTGGCCCCAAGATAT	344			
Oy	315	tggaaaagactatccaagaatcaagtcagtcagatgcatctctgtgtagcatctaaag	374			


```

Db      345  TGGCAACAACCTTCTCGAGAGATTAACTACTGAGANTGAGACCAGCGTGCGCTTAAG 404
Oy      375  cctgagaccacct 388
        11111 11111
Db      405  CCTGGAACACCTCT 418

RESULT  3
LOCUS   AE006493
DEFINITION Streptococcus pyogenes M1 GAS strain SF370, section 22 of 167 of
the complete genome.
VERSION AE006493 AE004092
KEYWORDS AE006493.1 GI:13621538
SOURCE  Streptococcus pyogenes M1 GAS.
ORGANISM Streptococcus pyogenes M1 GAS
Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 12091)
Ferretti,J.J., Moshan,W.M., Adjic,D., Savic,D., Savic,G., Lyon,K.,
Primeaux,C., Sezate,S.S., Surorov,A.N., Kenton,S., Lai,H., Lin,S.,
Qian,Y., Jia,H.G., Najaf,F.Z., Ren,Q., Zhu,H., Song,L., White,J.,
Yuan,X., Clifton,S.W., Roe,B.A. and McLaughlin,R.E.
Complete genome sequence of an M1 strain of Streptococcus pyogenes
Proc. Natl. Acad. Sci. U.S.A. 98 (8), 4658-4663 (2001)
2192684
PUBMED 11296296
TITLE   2 (bases 1 to 12091)
JOURNAL Ferretti,J.J., Moshan,W.M., Adjic,D., Savic,D., Savic,G., Lyon,K.,
MEDLINE Primeaux,C., Sezate,S.S., Surorov,A.N., Kenton,S., Lai,H., Lin,S.,
AUTHORS Qian,Y., Jia,H.G., Najaf,F.Z., Ren,Q., Zhu,H., Song,L., White,J.,
Yuan,X., Clifton,S.W., Roe,B.A. and McLaughlin,R.E.
Direct Submission
Submitted (10-APR-2001) Department of Microbiology and Immunology,
University of Oklahoma Health Sciences Center, 940 SL Young Blvd,
Oklahoma City, OK 73104, USA
FEATURES
source location/Qualifiers
1. 12091
/organism="Streptococcus pyogenes M1 GAS"
/strain="SF370"
/serotype="M1"
/db_xref="taxon:160490"
243..1115
/gene="SPY0263"
243..1115
/gene="SPY0263"
/note="Best Blastp hit = pir||A69879 conserved
hypothetical protein y100 - Bacillus subtilis
>g12337807|emb|CA74251.1| (Y13937) y100 protein
[Bacillus subtilis] >g12633950|emb|CAB13451.1| (Z99112)
similar to hypothetical proteins [Bacillus subtilis]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK3338.1"
/db_xref="GI:13621539"
/translation="MKGILKSLAGFYVESGOVYQTRARGNFRKRGFTPYVDIV
ESAEDNSGCIYLAHPRKRSIVRPVINDQAVIMSAREPFNSLDRFLILEHK
AHPVYVTSKMDLSPETIKRIGROYQAIIGYDFVSEELPLADKLTVMGQGV
GSLTLNRIAPELAIEIGISDSLGRGRTTAVSFYNTHGKIDTPESSLDIDA
NAEDNEAFPELRLSHCEKFRSCTHTHEPKCAVAALETGELMPVRYEHYQLFSEI
EKRRREYKVIKAK"
1125..1787
/gene="rpe"
/note="rpe"
/gene="SPY0264"
1125..1787
/gene="rpe"
/note="rpe"
/gene="rpe"
/BC_number="5.1.3.1"
/note="Best Blastp hit = dbj|BA06221.1| (AP001515)
unknown conserved protein [Bacillus halodurans]"
/codon_start=1
/transl_table=11

```

```

gene 1780..2412
CDS /gene="SPY0265"
1780..2412
/gene="SPY0265"
/note="Best Blastp hit = pir||CG9879 hypothetical protein
y105 - Bacillus subtilis >g12337809|emb|CA74253.1|
(Y13937) y105 protein [Bacillus subtilis]
>g12633952|emb|CAB13453.1| (Z99112) y105 [Bacillus
subtilis]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK3340.1"
/db_xref="GI:13621541"
/translation="MSKVALFAGDLSYISRDFYFVGIDRGSFLFLENGLPUNMAG
DDSVSOKAFDPIKEKAELEFITAHPKNDTDEIALKEVAREPPEAVTIFGAFGRM
DHLSTNIFPSPGTAIFPAQJALRQOQMITYRAGQLIHQBEGMTVAMABGEA
DLITGAKFELQDNEFFKRTYSSNAFHQPIVLSPLSGYLIILIOSKMS"
2414..3685
/gene="SPY0266"
2414..3685
/gene="SPY0266"
/note="Best Blastp hit = gb|AAFG3223.1|AE003892.8
(AE003892) conserved hypothetical protein [Xylella
fastidiosa]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK3341.1"
/db_xref="GI:13621542"
/translation="MDLILFLVLLGLGAYLLEKVNGLQHQLAQTFEGNADNSDQ
MYVOLDTPANKOOLELTOLMNRQAGLYEOLTDIRVLYHRSLSDSRDRDKLEKINO
OVNSOLTKMOESNEKLEKRMROIIVEKLEKTDRIHAFHSLSKQESVNGKLEMR
SVAODVGTLNKYLSNTRKRGIGELQIGOIIEIDIMSSQYEREPVYSGSSREVEAI
KLPNGCGGYIYIPIDSKRPLDGYRLDEPAVYCKLAIEASRKALLAIKRPANDH
KRYLNPETTNFVMPFLPTEGLKSEYVRNASEFDSLRRENTVYAGPSTLSALNSLS
VGFKTLINQKNADDISKILGNVLFEDKGGGLAKAQKQMTANNTLDLIDSTRINAI
VALNTVETVYQDQATKSLNMPLEEBENNEN"
3675..4613
/gene="cbf"
/note="cbf"
/gene="SPY0267"
3675..4613
/gene="cbf"
/note="cbf"
/note="Best Blastp hit = pir||G69818 CMP-binding factor
homolog yham - Bacillus subtilis
>g12226127|emb|CA74420.1| (Y14078) Hypothetical protein
[Bacillus subtilis] >g12633329|emb|CAB12833.1| (Z99109)
similar to CMP-binding factor [Bacillus subtilis]"
/codon_start=1
/transl_table=11
/product="putative CMP-binding factor"
/protein_id="AAK3342.1"
/db_xref="GI:13621543"
/translation="MKINQMKRDKQFEGEYLIKSAEVRKTRAGKDFISLTFODNGEI
SGNIMDAPQYVNEEFPAGKVVFMKGRREVYNGTPQVNOTTLNVRPGEKNDPKFEK
APVSATREVDYILEQMFKIENATWQIVATLRYDKDEEYTPAKTNRHARESGIAY
HTATVRLADSTIGDIPDLNKSILFAGMLMDLAKYIELTGPNDIETVYRGULIGHS
LINEITKVISLQIDTKEVIVLHVLVLSHGDLEGSSPVPRIMEAELIIMIDNI
DANMMMTALSRVSEGMTNRIIFADNRSFYKPNY"
4880..5719
/gene="pur"
/note="pur"
/gene="SPY0268"
4880..5719
/gene="pur"
/note="pur"
/note="Best Blastp hit = emb|CAA10902.1| (AJ222642) pur

```


--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

[illegible][illegible]

Db	682	AATTTGCAAGAGATGTGAACGTTTAGAATGTGCTGGAGCGTACCTTGTACATATTGC	623
Qy	129	atcatgatatgagcatlittgtccccaattactatattgacgtccagttatlyaaagtltg	188
Db	622	GTAATTTGACGGGACATTTTGTGGATTAATTAACCTTTTGGCGTGGGGCTTTTCGGGCACAT	563
Qy	189	agaagacacacaagaagcatatttgatgttcaacctatggttacaaatccctctgta	248
Db	562	CCCGTACAAACATCATTTATTTTCATGTTCACATGATGTGTGAANAATTCCTGAAAATAT	503
Qy	249	gttgaaccttggcaaaagctggtcctctggttttacaatttcaagtagagacatcaaa	308
Db	502	GTGGGAGATTTTGGCAAAAGCAGGACCTGATTACATGATGATTCATGTTGTAAACCACTCAT	443
Qy	309	gataactgaaagaactatccaagaatcaagtcacatgcatgcatgatcttctgtttaga	368
Db	442	CATATTCAACGGGACACTT---CAAAAATTAGAATGCCGCATGATAAGCTTCGGTAGTT	386
Qy	369	ttaaagccttggagcccccggt	388
Db	385	ATTAACTCTGGAACTCTCTG	366

RESULT	5
BC019126	
LOCUS	BC019126 1316 bp mRNA linear ROD 11-DEC-2001
DEFINITION	Mus musculus, clone MGC:29410 IMAGE:5008075, mRNA, complete cds.
ACCESSION	BC019126
VERSION	BC019126.1 GI:17512310
KEYWORDS	MGC.
SOURCE	house mouse,
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 1316)
AUTHORS	Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (07-DEC-2001) National Institutes of Health, Mammalian

REMARK	COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov	Contact: MGC help desk

```

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Plate: 39 Row: e Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13385145.
Location/Qualifiers
1..1316

```

```

tissue_type="Mammary tumor,
old mouse. Taken by biopsy."
/clone_lib="NCI_CGAP_Mam2"

```

```

CDS
    /lab_host="DH10B"
    /note="Vector: pCMV-SPORT6"
    58..744
        /codon_start=1
        /product="unknown (protein for MG:29410)"
        /protein_id="AAH19126.1"
        /db_xref="GI:17512311"
        /translation="MASGCKIGPSITLSDLANLGAECTRLMIDSGADYIHLDVDMGHPV
        PNITFGHPVYESLKRQLGQDFPFDMHMMVSRPEQWVYPMAYAGANTYTHLEATENG
        ALIIDIRENMKVGLAIRPGTVEYLA PMANQIDMALVMTVEPGFGQKEMEDMKPV
        HMLETPDLEIDVGGVGPDYVQKCAEAGANMIVSSAIIKRSDDPRAVYINLLRNVC
        EAAOKRSIDR"
BASE COUNT    392 a    253 c    314 g    357 t
ORIGIN

```

Query Match	Score	DB	Length
Best Local Similarity	24.9%	96.8	1316
Matches	211	Conservative	0
		Mismatches	142
		Indels	9
		Gaps	2

RESULT	6
LOCUS	BC006953
DEFINITION	BC006953 2357 bp mRNA linear ROD 12-JUL-2001
ACCESSION	BC006953
VERSION	BC006953.1 GI:13905310
KEYWORDS	MGC.
SOURCE	house mouse.
ORGANISM	Mus musculus.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 2357)
TITLE	Strausberg, R.
JOURNAL	Direct Submission
REMARK	Submitted (27-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: gcaps-remail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys CNA Library Preparation: Life Technologies, Inc. CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalobebcm.tmc.edu.
 Villalon, D.K., Luna, R.A., Hale, S.M., Huylk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>
 Series: IRAX Plate: 5 Row: f Column: 16
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

Source Location/Qualifiers
 1..2357
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="MGC:6936 IMAGE:2811899"
 /tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months old, gross tissue."
 /clone_lib="NCI CGAP_Mam5"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 9..239
 /codon_start=1
 /product="Unknown (protein for MGC:6936)"
 /protein_id="AAH06953.1"
 /db_xref="GI:13905311"
 /translation="MASGCKIPSPSTINSPLANIGAECLMLDLSGADYLDHLDVMDGHFV
 PNITGHPEVESLRFQLQGDPPEDMIMVSRP"
 BASE COUNT 753 a 414 c 490 g 700 t
 ORIGIN

Query Match 24.9%; Score 96.8; DB 10; Length 2357;
 Best Local Similarity 58.3%; Pred.No.3.3e-16;
 Matches 211; Conservative 0; Mismatches 142; Indels 9; Gaps 2;

33 aaatagctccttcgagtccttccttcgagcttcgcaattgctccgagctcagcgc 92
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 24 AAATGCGGCCGCTCCATCTCAACAGCGACCTGGCCAACTGGGGCCGAGTGCCTGCGC 83
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 93 atgtccactcgcgcgcgcgtatgctccacatgcatcatcatgagtgagcatttgc 152
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 84 ATCTGAGACTCGGGGCCGACCTGACCTGATGATGATGAGCGGCGCATTTTGTCC 143
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 153 aattcaactatggcgctccagttatgaaagttgagaagca-----cacaaaggca 206
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 144 AATATCACTTTTGTGACCTGCTGTGTAAGGCTCGAAGACGCTAGGCCAGACCT 203
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 207 tatttgatgtacaccttaagttacaatcctcttgatattgttaacaccttgcaaaa 266
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 204 TTCTTGATATGACATGAGGTGTCTAGGCCATACAGGTGGTAAACCAATGCGCTGTG 263
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 267 gctgtgctcttgctttatcatcctgtagagacatacaagaataagtaagaact 326
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 264 GCAGTGGCCCAATCAATATACCTTCACTCTGAG---GCACTGAGAAACCCAGGGGCTTTG 320
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 327 atccaaagaatcaagtcacatgcatgattcctgtgtgtagcatlaaagcctggagcccc 386
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 321 ATTAAAGACATCCGGAGAGATGGATGAAGTTGGCTTCCATCAACACGACGAACATACA 380
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 387 gt 388
 ||
 381 GT 382

RESULT 7
 LOCUS AP003004 338579 bp DNA linear BCT 15-MAY-2001
 DEFINITION Mesophilobium loti DNA, complete genome, section 11/21.

ACCESSION

AP003004 BA000012
 AP003004.2 GI:14024426

KEYWORDS

Mesophilobium loti (strain:MAFF30309) DNA.
 Mesophilobium loti
 Bacteria: Proteobacteria: alpha subdivision: Rhizobiaceae group;
 Phyllobacteriaceae: Mesophilobium.

AUTHORS

1 (sites)
 Kaneko,T., Nakamura,Y., Sato,S., Asamizu,E., Kato,T., Sasamoto,S.,
 Matanabe,A., Idesawa,K., Ishikawa,A., Kawashima,K., Kimura,T.,
 Kishida,Y., Kiyokawa,C., Kohara,M., Matsumoto,M., Matsuno,A.,
 Mochizuki,Y., Nakayama,S., Nakazaki,N., Shimpo,S., Sugimoto,M.,
 Tachibana,C., Yamada,M. and Tabata,S.
 Complete genome structure of the nitrogen-fixing symbiotic
 bacterium Mesophilobium loti
 DNA Res. 7 (6), 331-338 (2000)

TITLE

2 (bases 1 to 338579)
 Kaneko,T.
 Direct Submission
 Submitted (05-DEC-2000) Takakazu Kaneko, Kazusa DNA Research
 Institute, The First Laboratory for Plant Gene Research, Yana
 1532-3, Kisarazu, Chiba 292-0812, Japan
 (E-mail:kaneko@kazusa.or.jp,
 URL:<http://www.kazusa.or.jp/rhizobase/>,
 Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)
 On May 11, 2001 this sequence version replaced gi:11994979.

REFERENCE

1..338579
 /organism="Mesophilobium loti"
 /strain="MAFF303099"
 /db_xref="taxon:381"
 /complement(47..361)
 /gene="mll4353"
 /complement(47..361)
 /gene="mll4353"
 /note="hypothetical protein"
 /codon_start=1
 /transl_table=11
 /protein_id="BAB51030.1"
 /db_xref="GI:14024427"
 /translation="MRLRHMAAGFNAAGNIPPTLTALVNDGIYRRNRNPRYL
 GAPLYLGLGVAASFMTLVLPILMTVINTGVIAPERLERKKEFDATRAYRVRRL
 W"

FEATURES

Location/Qualifiers

source

gene

CDS

gene

980..1483
 /gene="mllr4354"
 980..1483
 /gene="mllr4354"
 /note="unknown protein"
 /codon_start=1
 /transl_table=11
 /protein_id="BAB51031.1"
 /db_xref="GI:14024428"
 /translation="MRGLTALVSSLYLVASAPVLAQAQTKIGOHNAKGTYSQASGG
 KVCYVLPIDKQPTLIDHGDMEFEVSQRPQGVSEPOFVAGVYNQESKATVITDK
 KSSMPTRGKSAVNAEAEPEVLIAMKGTDMKVTAKSGRNPYSVPSLKGISAL
 SSIAKCK"
 complement(1541..2884)
 /gene="mll4356"
 complement(1541..2884)
 /gene="mll4356"
 /transl_table=11
 /codon_start=1
 /transl_table=11
 /product="chloride channel protein"
 /protein_id="BAB51032.1"
 /db_xref="GI:14024429"

gene

complement(1541..2884)
 /gene="mll4356"
 complement(1541..2884)
 /gene="mll4356"
 /transl_table=11
 /codon_start=1
 /transl_table=11
 /product="chloride channel protein"
 /protein_id="BAB51032.1"
 /db_xref="GI:14024429"

CDS

AVLADKQALFHMIDGGGRREYPLATPLGFLVCAWLAARAFEPGSGSGIPQAI
 ARHLRDDDRSHLSLRILAAGIALITVGLFCGASIGREGPIYOVASIMLQAAKRG
 MAOARGLLIAGSAGIAAAMPNPLAGIVAIEMGRTYAPRTNGVLTAVIAGLASG
 GLGNTYFVGSGKETIAPADMPVIVGCGITGGREAFSLAIKTRIRIRWNOOP
 LRAILVAVCGIGAVIGIYASGGLTFEGGIGQAKRAVCGTLPMPFEKFLAGILS
 MWSGIPGGLFAPSLAVAGIGSSLLGLFASAGAAALLGMAGTFAAGVQAPMTAFVIT

QY	316	g9aaagaactatccaagaatcaagatcacatgagcagatgattccctggtgtagcatcaaacg	375
Db 321181	--GACCGCTCGCTGTGACAGCATTCAGGAAATCTTTGGCAGACAGAGCCCGGTGTGCTCAATC	321238	
QY	376	ctggagaccccg	387
Db 321239	CGCGACGCCG	321250	
RESULT	8		
LOCUS	AL596170/c	313450 bp	DNA linear BCT 04-DEC-2001
DEFINITION	Listeria innocua Clp11262 complete genome, segment 8/12.		
ACCESSION	AL596170		
VERSION	AL596170.1		
KEYWORDS	GI:16414292		
SOURCE	Listeria innocua.		
ORGANISM	Listeria innocua		
REFERENCE	Bacteria: Firmicutes: Bacillus/Clostridium group: Bacillus/Staphylococcus group: Listeria.		
AUTHORS	1 (sites)		
	Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloeker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouni, F., Cove, E., de Daruvar, A., Dehoux, P., Dommann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Ertlan, K. D., Fehli, H., Portillo, F. G., Garrido, P., Gautier, L., Geisel, M., Gomez-Lopez, N., Hain, T., Haut, J., Jackson, D., Jones, L. M., Kaerst, U., Kreft, U., Kuhn, M., Kunst, F., Kurapat, G., Madueno, E., Maitouram, A., Vicente, J. M., Ng, E., Nedjati, H., Nordisiek, G., Novella, S. de Pablo, B., Perez-Diaz, J. C., Purrnell, R., Remmel, B., Rose, M., Schueter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J. A., Voss, H., Wehland, J. and Cossart, P.		
TITLE	Comparative genomics of Listeria species		
JOURNAL	Science 294 (5343), 849-852 (2001)		
MEDLINE	21537279		
REFERENCE	2 (bases 1 to 313450)		
AUTHORS	Glaser, P., Frangeul, L. and Rusniok, C.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-JUL-2001) Glaser P., Institut Pasteur, Genomique des Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE		
COMMENT	E-mail: palaser@pasteur.fr Phone: +33 (0)1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.		
FEATURES	Location/Qualifiers		
source	1..313450		
gene	/organism="Listeria innocua"		
CDS	/strain="Clp11262"		
	/db_xref="taxon:1642"		
	complement(287..1444)		
	/gene="lin1790"		
	complement(287..1444)		
	/gene="lin1790"		
	/note="similar to transmembrane transport proteins"		
	/codon_start=1		
	/protein_table=1		
	/protein_id="CAC97021.1"		
	/db_xref="GI:16414293"		
	/translation="MGAKGRFMTITVVAISGLQVLLPLIALILEEKVSAGINIEF HANGIVGLVLPSEPIAPLAKYGRKILIVGGGLVAIAIPAFIMNLEWFLRLIL IGIDGHLHRSOTIGAMSDPSKRGNNMAYIGFSGVRAIGQIVNLAKVANNLPEF FLSGILVLIAMSILWFIIRNDPVEKAKIRIKISFPGSIKRSQYKAMVAMIPFLYG ILEGLNATFPVVGRLDGLDTMAMLIISFSGVTIIFQVPIGISKFGDKMLPL TGAAGVAVLTAETKIPFIVFIEFFILGILGSLYSGLSYMDDLPLELPAENIIV GMFESLGIIGPSATGMGMIIFGNQIPYFVAGILLCCLLALGKREAH"		
	complement(1508..1579)		
	/note="tRNA scan -SE vs 1.3 result - Cove score = 69.21"		
	/product="tRNA-Glu"		
	complement(1585..1675)		
	/note="tRNA scan -SE vs 1.3 result - Cove score = 63.76"		

tRNA	/product="tRNA-Ser" complement(1712.. .1784) /note="CRNScan-SE vs 1.3 result - Cove score = 85.46"
tRNA	/Product="tRNA-Asn" complement(1791.. .1864) /note="CRNScan-SE vs 1.3 result - Cove score = 92.91"
tRNA	/product="tRNA-Ile" complement(1891.. .1961) /note="CRNScan-SE vs 1.3 result - Cove score = 73.1"
tRNA	/product="tRNA-Gly" complement(1984.. .2056) /note="CRNScan-SE vs 1.3 result - Cove score = 72.11"
tRNA	/product="tRNA-His" complement(2083.. .2155) /note="CRNScan-SE vs 1.3 result - Cove score = 79.39"
tRNA	/product="tRNA-Phe" complement(2160.. .2235) /note="CRNScan-SE vs 1.3 result - Cove score = 73.32"
tRNA	/product="tRNA-ASP" complement(2240.. .2313) /note="CRNScan-SE vs 1.3 result - Cove score = 78.6"
tRNA	/product="tRNA-Met" complement(2356.. .2445) /note="CRNScan-SE vs 1.3 result - Cove score = 66.89"
tRNA	/product="tRNA-Ser" complement(2468.. .2541) /note="CRNScan-SE vs 1.3 result - Cove score = 90.88"
tRNA	/product="tRNA-Met" complement(2553.. .2626) /note="CRNScan-SE vs 1.3 result - Cove score = 77.07"
tRNA	/product="tRNA-Met" complement(2681.. .2753) /note="CRNScan-SE vs 1.3 result - Cove score = 84.64"
tRNA	/product="tRNA-Ala" complement(2773.. .2846) /note="CRNScan-SE vs 1.3 result - Cove score = 86.85"
tRNA	/product="tRNA-Pro complement(2857.. .2930) /note="CRNScan-SE vs 1.3 result - Cove score = 74.21"
tRNA	/product="tRNA-Arg" complement(2940.. .3025) /note="CRNScan-SE vs 1.3 result - Cove score = 66.88"
tRNA	/product="tRNA-Leu" complement(3040.. .3114) /note="CRNScan-SE vs 1.3 result - Cove score = 87.88"
tRNA	/product="tRNA-Gly" complement(3130.. .3211) /note="CRNScan-SE vs 1.3 result - Cove score = 69.82"
tRNA	/product="tRNA-Leu" complement(3217.. .3289) /note="CRNScan-SE vs 1.3 result - Cove score = 88.37"
tRNA	/product="tRNA-Tyr" complement(3330.. .3405) /note="CRNScan-SE vs 1.3 result - Cove score = 95.38"
tRNA	/product="tRNA-Thr" complement(3414.. .3486) /note="CRNScan-SE vs 1.3 result - Cove score = 86.53"
tRNA	/product="tRNA-Val" complement(3503.. .3612)
gene	/gene="tRNA-S5" complement(3503.. .3612)
rRNA	/gene="tRNA-S5" complement(3693.. .6623)
gene	/gene="tRNA-23e" complement(3693.. .6623)
rRNA	/gene="tRNA-23s" complement(3693.. .8423)
gene	/gene="tRNA-16s" complement(6869.. .8423)
terminator	/gene="tRNA-16s" complement(8850.. .8868) /gene="lin1791"

gene
CDS
/gene="lin1791"
complement(876, .9322)
/gene="lin1791"
/note="similar to transcription regulators (Fur family),
perr in B. subtilis"
/codon_start=1
/transl_table=1
/protein_id="CAC97022.1"
/db_xref="GI:16414294"
/translation="MAVSNAATLKEAVDVLLKKTGVRITPORHALLFLINSHPTTAD
IYRSLEGNPMVSATYVNNLRVFRDQDILKELSYGASSPDESTSNHIAITCNVCG
KIVDEHYPGIDVEHPFAAHVGYETIDNHRLEVGTCECEKOSNN"
complement(9350, .9353)
/gene="lin1791"
complement(9405, .10361)
/gene="lin1792"
complement(9405, .10346)
/gene="lin1792"
/note="similar to glycerate dehydrogenases"
/codon_start=1
/transl_table=1
/protein_id="CAC97023.1"
/db_xref="GI:16414295"
/translation="MKYLFITLDVPERIKNOIAKEFPDDTTFPSIDHFTNUGENVIT
TYGNTIEKIKKATNTKLIVTSAGVDSLPTKITQKIKVANVRGIIHAPGEYAL
SEMLSHVKKAAFFYKMQEKEMASEEPITELAGTIVVAGTGAIAKAAVEAQAQFME
VIGINTGHPYKPSKTYAMTLEKVAIPADFPVSPOTEETSGIYOLSFPEKMTN
AVFINIGRSAVELETLEKASEQIAHFVLDVPEPPLPAESYLMQASVTTTPHS
GHSIDYDRSEIFEFENINHLKENTKRLNEIDLKNGY"
complement(10356, .10361)
/gene="lin1792"
10476, .11787
/gene="gsab"
10476, .10481
/gene="gsab"
10489, .11787
/gene="gsab"
/codon_start=1
/transl_table=1
/product="glutamate-1-semialdehyde aminotransferase"
/protein_id="CAC97024.1"
/db_xref="GI:16414296"
/translation="MDHSMSKLLHDEALLHIVGVNSPSRNSKGVGGTIPVTERANG
AIFYVDGNKYIDYIAAFGLITIGHAHPHTTEATTKAONGVYXGPTTKETIPAKML
KEAIPSEKRYPTNSGTEANMTTIRARATGDKITKAGCYHGHFDLVVAGSGP
STIGIPDSAGVTSSTAEVITVPPNDLTSKELAVMGDVAIVLEPITVGNFGWAP
AEGLEAVNELAHNGLSLVYDEVITAIFRMYGAAQNYLGVIDLTAMGKIIIGGGPI
GAYGGRVDIMEKVAPLGPAYOAGTHAGNPASISGLACLEVLEEGILYDFEKGSM
KEGIERKAIKHGIATVNOYVAGLTYFTDEPTVNAEAGATNGDLFGRFKMLEEG
INLAPSKYEMAFITSAHSEADILETTIOAVDTVHGKLVQCK"
11879, .11884
/gene="lin1794"
Join(11879, .12981, 12990, .13016)
/gene="lin1794"
11893, .12981
/gene="lin1794"
/note="similar to hypothetical proteins"
/codon_start=1
/transl_table=1
/protein_id="CAC97025.1"
/db_xref="GI:16414297"
/translation="MKFGARILKTGIAITLALFIAOLCNSPSPSLAGISAVFALOPSI
YRSYRTILERAQGNVGIATITFEGYIGNDFLIGVASTICVALLMORLENTTGLA
VYTLITVDSPGNDLEIALIRGCTIMGLAAFTYNLFFLPKTEVSFLQLYNTNS
EIVRWATKLNRHADPFLKKDMEMQKDLNRTNLYRREERFFLKRNAISKGRKI
AVRKQMLCSQKGFELKIQHRENDYDLPKQKOLIRQHIDYLTDKHKOULLYID
KVSIDLEYVSHLAQDPODMLFLREMEETKDEYEDMMDKYHLMRIIASVFAVQET
IDYLEKLIHSFKIRHTEENOIDINVEE"
complement(Join(12990, .13016, 13022, .13549, 13557, .13562))

Query Match 23.6%; Score 91.4; DB 1; Length 313450;
Best Local Similarity 56.08; Pred. No. 2,4e-14;
Matches 173; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

Qy 3 agaaagaagaagaatgaggatgacacccgaataagctctctcgatgctctctccgac 62
Db 157025 AAACGAGAAAGCCGAGATTAAAAATGAGAAATGACTCTCTCGATTTTAAGTGCAGAC 156966
Qy 63 ttcgccaattgctccgaagctcaagcgacatgctccacttcgcgccgagctgctcac 122
Db 156965 TTTCGAATATCTTGCAGAGATATTAAAGCAATACAAAATTACGGCGGATTTATTCAT 156906
Qy 123 atggacatcatgagctggagcatctgctcccaatttaactatlgcgctccagttatgaa 182
Db 156905 GTTGATGTAATGACGCGCATTTCTTCGCAAAATATCACTTTCGCTCCGCTGCTGTA 156846
Qy 183 agttggaaagacacaaagcatttgatgctacatttggttaacaaatccctct 242
Db 156845 GCGATTCTCCGGAACAAACTACCACTTGATGTTATTTAATGATTAATCCGAT 156786
Qy 243 gattatgttgaaaccttgcaaaagctgctgctctggtttacattccagtagagaca 302
Db 156785 ACATACATCCCTGAAATTTCGAAAGGACGGCGGATTTATTCACGTTTATGTAAGCG 156726
Qy 303 tcaaaagat 311
Db 156725 TGCACGCAT 156717
RESULT 9
LOCUS AL591981/c 347050 bp DNA linear BCT 04-DEC-2001
DEFINITION Listeria monocytogenes strain EGD, complete genome, segment 9/12.
ACCESSION AL591981 AL591824
VERSION AL591981.1 GI:16411141
KEYWORDS
SOURCE Listeria monocytogenes.
ORGANISM Listeria monocytogenes.
REFERENCE
AUTHORS Glasner,P., Frangeul,L., Buchrieser,C., Rusnok,C., Amend,A.,
Bachner,F., Berche,P., Biocker,U., Brandt,P., Chakraborty,T.,
Charbit,A., Chetoui,A., Couve,E., de Daruvar,A., Dehoux,P.,
Domann,E., Dominguez-Bernal,G., Duchaud,E., Durant,L.,
Dussurget,O., Entian,K.D., Fsihl,H., Portillo,F.G., Garrido,P.,
Gautier,L., Gobeil,W., Gomez-Lopez,N., Hain,T., Hauf,J.,
Jackson,D., Jones,L.M., Kaserst,U., Kreft,J., Kuhn,M., Kunst,F.,
Kurapkai,G., Maduno,E., Maitournam,A., Vicente,J.M., Ng,E.,
Nedjari,H., Nordstedt,G., Novella,S., de Pablo,B., Perez-Diaz,J.C.,
Purcell,R., Rammel,B., Rose,M., Schlueter,T., Simoes,N.,
Tierrez,A., Vazquez-Boland,J.A., Voss,H., Wehland,J. and Cossart,P.
Comparative genomics of Listeria species
Science 294 (5543), 849-852 (2001)
TITLE JOURNAL
MEDLINE 2 (bases 1 to 347050)
REFERENCE 21537279
AUTHORS Glasner,P., Frangeul,L. and Rusnok,C.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-2001) Glasner P., Institut Pasteur, Genomique des
Microorganismes Pathogenes, 25 rue du Docteur Roux, 75724 Paris
Cedex 15, FRANCE
COMMENT E-mail: pglaser@pasteur.fr
Phone: +33 1 45 68 89 96, Fax: +33 (0)1 45 68 87 86.
FEATURES
source
1. 347050
/organism="Listeria monocytogenes"
/strain="EGD-e"
/db_xref="taxon:1639"
complement(95, .841)
/gene="lmo1688"
complement(95, .841)
/gene="lmo1688"
/note="similar to glucose 1-dehydrogenase"
/codon_start=1
/transl_table=1
/protein_id="CAC97066.1"


```
/note="similar to unknown proteins"
/codon_start=1
/transl_table=11
/protein_id="CAC99774.1"
/db_xref="GI:16411150"
/translation="MOVOKRFEVLLPALVLYGISLAFQGNADLNTIANSCLL
FLAVLILTMARLNKIDIMFCEVILYIILHLVYIAVGVNSTYGNHIOKE
MINIPPTIENFOOTLPMPPTIIOIGVNLCLPLSFMLFKITSTAGTKLLVIF
LTSGLIELQFAOTMTMGFESISLPKRSTDDIDILNLSIGLILVYALPSRK

Query Match
Best Local Similarity 54.08; Pred. No. 5,2e-14;
Matches 207; Conservative 0; Mismatches 173; Indels 3; Gaps 1;

QY 3 agaaagaaagaaagatggaatgacacgaataaataagctcctcgtcctcctcgcac 62
Db 139624 AACAGAAAGCCGAGGTATTAAAGATGGGAAATAGCTCCTTGATTTTAAGTCAGAC 139565
QY 63 ttgcgcaattggctccgaggtccagcgatcgtccacatcgcgcgcgcgcgcgcgcac 122
Db 139564 TTTCGACCTTTGCAAGAGATTAAAGATAGAAATTCGGGTGCGAGATTACATCCAT 139505
QY 123 atggacatcatggaatggcattcgtcccaatttaactatggtcgcgtccagatgaa 182
Db 139504 ATTGACGTTATGANTGACATTTGTTCCAAATATTACATTGGACCTGCTGTGTCCA 139445
QY 183 agttgagaagacacacaaagcatattgagatgcatcactatggtatcaaatccctc 242
Db 139444 GCCATTGCTCCAGAAAGCAATCCACTTGATGTTCATTAAATGAAATCCGGAT 139385
QY 243 gattatgtgaaacctggcgaagctggtcgtcgtcgttataattcaatgagagaca 302
Db 139384 ACATACATTCAGAGTTTGCAAAAGCGGTCGACATACATACCGTCGATGCGAAGCT 139325
QY 303 tcaaaagataactggaagaactatccaaagaatcaagtcacatgcatgcatcctggt 362
Db 139324 TGTAACGAT--CTTCAATCGACATACATTAATTCGTTCTTATGTTGGAAGCCGGA 139268
QY 363 gtacatctaaagcctggagacc 385
Db 139267 GCACTACTTAACCGACGACACC 139245

RESULT 10
AX085734 1205 bp DNA linear PAT 09-MAR-2001
LOCUS AX085734
DEFINITION Sequence 12 from Patent WO0112790.
ACCESSION AX085734
VERSION AX085734.1 GI:13275684
KEYWORDS
SOURCE
ORGANISM Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1205)
AUTHORS Bandman,O., Lu,D.A., Yue,H., Tran,B., Hillman,J.L., Baughn,M.R.,
Lal,P. and Tang,Y.T.
Isomerase proteins
Patent: WO 0112790-A 12-22-FEB-2001;
JOURNAL Incyte Genomics, Inc. (US)
FEATURES
source
1..1205
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 2481256CB1"

BASE COUNT 331 a 224 c 291 g 359 t

Query Match
Best Local Similarity 52.9%; Score 88.8; DB 6; Length 1205;
Matches 206; Conservative 0; Mismatches 147; Indels 9; Gaps 2;
```

```
QY 33 aaatagctcttcgatgctctcttcgcgaacttcgccaatttgcttcgaggtcagcg 92
Db 21 AAGATTGGCCCGTCCATCTCAACAGCGACCTGCGCAATTAAGGGCGGAGCTCCCG 80
QY 93 atgctcacttcggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 152
Db 81 ATGTTAAGACCTGGGGCGGATATCTACCTGAGCTTAATGAGACGGCATTTGTTC 140
QY 153 aatttaactatggtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 206
Db 141 AACATCACTTTGGTACCTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 200
QY 207 tatttgatgtaacacctatgtaacacctatgtaacacctatgtaacacctatgtaacacct 266
Db 201 TTCTTTACATGCAATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 260
QY 267 gctggtccttcgtttatcattcctcgtagagacacaaagataactcgtgaagaact 326
Db 261 GCAGGAGCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 317
QY 327 atccaagaatcaagtcacatgcatgcatgcatgcatgcatgcatgcatgcatgcatgcat 386
Db 318 ATTAAAGACATTCGGGAGATGGATGAAGTTGGCTTGCCATCAACAGGAACCTCA 377
QY 387 gt 388
Db 378 GT 379

RESULT 11
BC016764 1778 bp mRNA linear PRI 09-NOV-2001
LOCUS BC016764
DEFINITION Homo sapiens, clone MGC:22637 IMAGE:4107315, mRNA, complete cds.
ACCESSION BC016764
VERSION BC016764.1 GI:16876981
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1778)
AUTHORS Strausberg,R.
JOURNAL Direct Submission
Submitted (05-NOV-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-tr@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
Contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 32 Row: C Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
location/Qualifiers
1..1778
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:22637 IMAGE:4107315"
/rclone_type="Bone marrow, acute myelogenous leukemia"
/clone_id="NIH_MGC_55"
/lab_host="DH10B"
```


gene
CDS
/gene="BME11114"
2647 . 3675
/EC_number="1.1.1.1"
/codon_start=1
/transl_table=11
/product="ALCOHOL DEHYDROGENASE"
/protein_id="AAL52295.1"
/db_xref="GI:17983086"
/translation="MRALQLDLDRRLLEITDIAPPLPGMGVTVRIKAVLNHIDVWGSRMGFAFRKRPMLYVIGASGVVDVAVGCVSNLLPGOLYSIGARTGGLCACEGRDNLCEHVGAVGHGDLDFGACFCAVNLPAFLVAPPGDVAIGAAPVPTGVEHMLPNNAKLEGSEVIVDAGSGSGTGAIOIAKMGCTVTTGSSNKKIKAKALGDHYINRREDEGCVIRKLTKKKGVDFVEHVADTWACSMLOMKRGALVTGSGTSGVSTMMNLMLQLFDQOLITFSSFGCRMENMADAKMARGIHVPIIDTVVGSIDITALKMGEDRDFVGIILQID"
3692 . 4615
/gene="BME11115"
3692 . 4615
/gene="BME11115"
/EC_number="2.3.1.-"
/codon_start=1
/transl_table=11
/product="LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE"
/protein_id="AAL52296.1"
/db_xref="GI:17983087"
/translation="MFKLKLLFRNSRKLKQFNWYMLMAQAVFVLGLGRLTPAKAASISFSARVARIIGLTPFRHKVATDNLKRAYEKSEAEIVKJARMBDSMARLFAEYIFLDDAVEDPVPVAKGVGLVEVEGIPLEEGLRDCKRHPFTTAHTGNFELLPCATPFGIVNTPRLFRNNPVYINKYLYKAKARTMGHLYPKKAGANAAAGLIEDGGNVGMLYQKRSKRVPSFPNNRYKTNPLAKRLAROYDCDYTPARCTIRLPGGRILETIERMELPDESGDILINATYQLNDVNEOVNREYPGQMMWFHRWG"
4632 . 5450
/gene="BME11116"
4632 . 5450
/gene="BME11116"
/EC_number="5.1.3.1"
/codon_start=1
/transl_table=11
/product="RIBULOSE-3-PHOSPHATE 3-EPIMERASE"
/protein_id="AAL52297.1"
/db_xref="GI:17983088"
/translation="METGFWIKTRASKLIERFQRFRRNRNSRIFAPVPIRASHAKEITAMPILAIABISLADPARLKEEDVAVLEAGADWVHLIDWDGHPVNPITFGPVVYVARIIPRKAFEDFADHMLIAPCDPYLPEPPAEGCDLITVHAEGPHTRSLDILALCKKAGALALNAPFAEALNTVNLDDILVWTVNPGFGGQKFTAPMLDKIRVRREMGDDPIDIDEVGGITPEPTGSGVAAACANLIVAGSAVYKGSNVSERYANIDALRAAAEARNR"
5477 . 6778
/gene="BME11117"
5477 . 6778
/gene="BME11117"
/EC_number="4.3.2.2"
/codon_start=1
/transl_table=11
/product="ADENYLOSUCINATE LYASE"
/protein_id="AAL52298.1"
/db_xref="GI:17983089"
/translation="MIPYSRPEMVAIWSPEKTRIMEPEIAYACALAEIGVYPKCAARTLTIKESVAKPVARDETLEATYKHDVIAFLNHLEFGTSDSPRYHOGMTSSDYDITTLNQLVRAADLLADLDRLVLEALAKRFAFKTKTIVRIGRSHGTHAEPITGLTFAFRTYAESKRNRARLVAAAEIATGAISAGVTEFANIDPVEEYCGAKGLEAEVRSQVILPDRHMAFPATGLVYIASSMENTAIEIRHMQRETVLEAEFFSPGQSSAMHKSNPVLPTENLGLARLVKMSYTPAMENVALMHNERDISHSVERAIGDPTITLDFALNLAGVIVLEKVIYPNMLKKNMKRPGVIGHSORVILLDQAGVSREDAIRLVQRNMKKWBGSDVPLEELADIEVAALASEBDIRREKFDLGTYTKTIVDIIKRYFG"
complement(7004 . 7879)
/gene="BME11118"
complement(7004 . 7879)
/gene="BME11118"
/EC_number="1.13.11.39"
/codon_start=1
/transl_table=11
/product="BIPEHENYL-2,3-DIOL 1,2-DIOXYGENASE III"

[illegible]

REFERENCE AUTHORS TITLE	Bacillus/Staphylococcus group; Bacillus. 1 (sites) Takami,H. and Horikoshi,K. Reidentification of facultatively alkaliphilic Bacillus sp. C-125 to Bacillus halodurans
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	Biosci. Biotechnol. Biochem. 63, 943-945 (1999) 2 (sites) Takami,H., Takaki,Y., Nakasone,K., Sakiyama,T., Maeno,G., Sasaki,R., Hirama,C., Fujii,F. and Masui,N. Genetic analysis of the chromosome of alkaliphilic Bacillus halodurans C-125
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	Extremophiles 3 (3), 227-233 (1999) 99411980 3 (sites) Takami,H., Nakasone,K., Hirama,C., Takaki,Y., Masui,N., Fujii,F., Nakamura,Y. and Inoue,A. An improved physical and genetic map of the genome of alkaliphilic Bacillus sp. C-125
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	Extremophiles 3 (1), 21-28 (1999) 99104645 4 (sites) Takami,H., Masui,N., Nakasone,K. and Horikoshi,K. Replication origin region of the chromosome of alkaliphilic Bacillus halodurans C-125
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	Biosci. Biotechnol. Biochem. 63 (6), 1134-1137 (1999) 99367711 5 (sites) Takami,H., Takaki,Y., Nakasone,K., Hirama,C., Inoue,A. and Horikoshi,K. Sequence analysis of a 32-kb region including the major ribosomal protein gene clusters from alkaliphilic Bacillus sp. strain C-125
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	Biosci. Biotechnol. Biochem. 63 (2), 452-455 (1999) 99209008 6 (sites) Takami,H. Genome analysis of facultatively alkaliphilic Bacillus halodurans C-125
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	Extremophiles in deep-sea environments, 249-284 (1999) 7 (sites) Takami,H., Nakasone,K., Ogasawara,N., Hirama,C., Nakamura,Y., Masui,N., Fujii,F., Takaki,Y., Inoue,A. and Horikoshi,K. Sequencing of three lambda clones from the genome of alkaliphilic Bacillus sp. strain C-125
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	Extremophiles 3 (1), 29-34 (1999) 99184646 8 (sites) Takami,H. and Horikoshi,K. Analysis of the genome of an alkaliphilic Bacillus strain from an Industrial point of view
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	Extremophiles 4 (2), 99-108 (2000) 20263314 9 (sites) Nakasone,K., Masui,N., Takaki,Y., Sasaki,R., Maeno,G., Sakiyama,T., Hirama,C., Fujii,F. and Takami,H. Characterization and comparative study of the rrr operons of alkaliphilic Bacillus halodurans C-125
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	Extremophiles 4 (4), 209-214 (2000) 20426005 10 (sites) Takami,H., Nakasone,K., Takaki,Y., Maeno,G., Sasaki,R., Masui,N., Fujii,F., Hirama,C., Nakamura,Y., Ogasawara,N., Kuhara,S. and Horikoshi,K. Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	Nucleic Acids Res. 28 (21), 4317-4331 (2000) 20512582 11 (bases 1 to 303249) Takami,H. and Takaki,Y. Direct Submission Submitted (22-MAR-2000) Hideto Takami, Japan Marine Science and Technology Center, Deep-sea Microorganisms Research Group, 2-15 Natsushima, Yokosuka, Kanagawa 237-0061, Japan (E-mail:takami@jamstec.go.jp,

COMMENT FEATURES SOURCE	URL:http://www.jamstec.go.jp/jamstec-e/bio/DEEPSTAR/Fresearch.html, Tel:81-468-67-3895, Fax:81-468-66-6364) On Jan 31, 2001 this sequence version replaced gi:10174886. Location/Qualifiers 1. 303249 /organism="Bacillus halodurans" /db_xref="taxon:86655" /note="alkaliphile" 104. .1102 /gene="BH2267" 104. .1102 /gene="BH2267" /codon_start=1 /transl_table=1 /product="penicillin-binding protein" /protein_id="BAB05986.1" /db_xref="GI:10174887" /translation="MKNHLTMEDKISTGATLMKKGEDIPRTASGFANRKRIPNE HHTREGISGCKLPTATQIYEAACKSFDTPLSWLDAPPNVTIHLTLTSGVP DYDEETDPEDFEDKMDVPTIHLRLDFLPKQHPMKFPQGRHYNNAGFILGL VSSGSVTRQDEVEANVEQRAGMHESGFAPEDLPKATGALGIDLEDGSMNTLYSL PVIGSDGKAYVAEDMKLMLMHHEILNETYTQKLTLPVHCEDDGYCYGVAWK QDGAISKYHVMGYDPVCFHSAFPTSNIGIYVVCANQSSGAVDVAALAEALFSEA" complement(1146. .1469) /gene="BH2268" complement(1146. .1469) /gene="BH2268" /note="BH2268 unknown" /codon_start=1 /transl_table=1 /protein_id="BAB05987.1" /db_xref="GI:10174888" /translation="MKVITLDCSTGRKFPQDASAVTLKGHVLVSGFFQSGEIEI TEQVVMFENHRRKTEMSDELIVDQVYIGRSTRKEIEFATSKGMPIRYSESELA STYES" complement(1636. .2373) /gene="BH2269" complement(1636. .2373) /gene="BH2269" /note="BH2269 unknown" /codon_start=1 /transl_table=1 /protein_id="BAB05988.1" /db_xref="GI:10174889" /translation="MKKYGVLGVFAISIGLMACNGOEQEBVEKNIDGASGETPST EETIDINLIPDVATVNGESIDKETVYVLEQMTMAQYGISLEGESKQMSILIE OTTEOLINEQLLOQAATERKEISEEIDELAFVAQFSSEALMEALLEGSSMDR LREIEHYRQGRKYVEETEAINTVEETIQARYEBEKQYTEDELPTPEELQNSINEQR LIAKEEQLEVLFEKLRDEGDITVHI" complement(2582. .3367) /gene="BH2270" complement(2582. .3367) /gene="BH2270" /codon_start=1 /transl_table=1 /product="ABC transporter (ATP-binding protein)" /protein_id="BAB05989.1" /db_xref="GI:10174890" /translation="MADILQVEYTLRRKAPRIIQISMOVQGEHMAIYGLNGSKT SLKLTITVEWTEGSHVAVLGNRYGHVPIQVRRIRIGVNSLSDRRHTKGGDTVEI VLSGHGVGVYVEIGENDVAKAAEALETFTLTLPTDERFLTISQGRKRAFLARAV AKPELFIIDPEPTTGILDLARQLQTTESVATYTGAPTEFLVTHTYPEIDPPIITHVM LKDGHVIAQGRKDEVLTKEMTTAAFGCSLDVIAQSGRYWKPM" complement(3378. .4310) /gene="BH2271" complement(3378. .4310) /gene="BH2271" /codon_start=1 /transl_table=1 /product="dipeptidase" /protein_id="BAB05990.1"
-------------------------------	---

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 03:38:53 ; Search time 6582.52 Seconds
(without alignments)
795.564 Million cell updates/sec

Title: US-09-300-482-298

Perfect score: 388
Sequence: 1 ggagaaagaagaagaatg.....ttaagcctggaccacct 388

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database :

EST :
1: em_estba:*
2: em_estlum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_lmv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	370	95.4	558	10	BF598003 sv02h10.y
2	365.4	94.2	418	9	AI437973 sa34c12.y
3	352.8	90.9	575	10	BM309659 sak65b08.
4	338.2	87.2	374	10	BG316079 sab65g09.
5	320.8	82.7	403	9	AV412302
6	310.4	80.0	405	9	AV412302 AV412302
7	307.8	79.3	574	10	AI442690 sa85b04.y
8	307.8	79.3	641	9	BI308956 EST530366
9	307.8	79.3	654	10	AM682881 NF001C10L
10	307.8	79.3	683	10	BG455207 NF102G10P
11	296.8	76.5	537	10	BF645593 NF016H02E
12	296.8	76.5	613	10	BG451782 NF094D04D
13	293.2	75.6	666	10	BE325053 NF119G01S
14	285.4	73.6	736	9	BF634692 NF064B05D
15	281.2	72.5	399	9	AM348508 GM210002B
16	281.2	72.5	448	9	AV425192 AV425192
17	269.8	69.5	467	10	AL371120 M8A42A06
					BI124136 I036P65P

18	260	67.0	745	10	BG526117	BG526117 56-78 Ste
19	250.4	64.5	506	9	AJ412644	AJ412644 AJ412644
20	247	63.7	571	10	BC907768	BC907768 Talr1162F
21	244.6	63.0	613	10	BF261410	BF261410 HV.CEAO00
22	244.6	63.0	816	9	AM983286	AM983286 HVSMEG001
23	244.6	63.0	837	10	BF254937	BF254937 HVSMEF000
24	244.6	63.0	838	10	BI951621	BI951621 HVSMEH000
25	243.8	62.8	659	10	BC608149	BC608149 WHE2497.E
26	243.8	62.6	488	10	BC606577	BC606577 WHE2958.B
27	243	62.6	494	10	BE404347	BE404347 WHE0441.B
28	242	62.4	506	10	BE601743	BE601743 HVSMEH009
29	241.4	62.2	268	10	BI993749	BI993749 sag92d02.
30	239.8	61.8	417	9	AU166524	AU166524 AU166524
31	238.2	61.4	526	10	BE94864	BE94864 WHE1258.D
32	237.4	61.2	407	9	AU069503	AU069503 AU069503
33	236.6	61.0	454	10	BF483212	BF483212 WHE1787.B
34	234	60.3	611	10	BM112847	BM112847 EST560383
35	232.4	59.9	669	10	BC599823	BC599823 EST504718
36	232.4	59.9	679	10	BC599735	BC599735 EST504630
37	232.4	59.9	699	10	BI432299	BI432299 EST535060
38	228.8	59.0	694	9	AU164966	AU164966 AV554048
39	226.8	58.5	483	9	AV554048	AV554048 AV554048
40	226	58.2	497	9	AI490023	AI490023 EST248362
41	226	58.2	552	9	AI772355	AI772355 EST253455
42	226	58.2	593	9	AM031337	AM031337 EST274791
43	226	58.2	630	9	AM223738	AM223738 EST300549
44	226	58.2	653	9	AI896020	AI896020 EST265463
45	226	58.2	729	10	BI928671	BI928671 EST548560

ALIGNMENTS

RESULT 1
LOCUS BF598003
DEFINITION sv02h10.y1 Gm-cl056 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl056-764 5' similar to TR:Q9SE42 Q9SE42 D-RIBULOSE-5-PHOSPHATE 3-EPIMERASE ;, mRNA sequence.

ACCESSION BF598003
VERSION BF598003.1 GI:11690327
KEYWORDS EST.

SOURCE soybean.
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 558)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Retsing,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

FEATURES
source 1..558
/organism="Glycine max"
/db_xref="taxon:3847"


```

/clone="GENOME SYSTEMS CLONE ID: Gm-cl056-764"
/clone_lib="Gm-cl056"
/issue_type="Whole seedling, 4 day old"
/lab_host="DH10B"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
from 4 day old seedling of P146916. The seedlings were
germinated in a growth chamber using germination paper.
Complementary DNA was synthesized from mRNA using a primer
consisting of a poly(dT) sequence with a XhoI restriction
site. EcoRI adapters were ligated to the blunt-ended cDNA
fragments followed by XhoI digestion. The cDNA fragments
were directionally cloned into the EcoRI-XhoI restriction
site of the pBluescript vector. The ligated cDNA fragments
were transformed into DH10B host cells (GibcoBRL). This
library was constructed in the laboratory of Dr. Randy
Shoemaker."

```

```

BASE COUNT      150 a      113 c      134 g      159 t      2 others
ORIGIN

```

```

Query Match      95.4%; Score 370; DB 10; Length 558;
Best Local Similarity 97.1%; Pred. No. 4e-96;
Matches 376; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

OY 3 aagaagaagaagaatggaatgacacccgaaataagctccttgatgctcttcgcgac 62
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 49 AGAAGAGAGAGAGAGATGGAGTGAACCGAAATTTGCTTCATGCTCTTCGAC 108
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 63 ttgcgaattggtctccgaggtcagcgatgctccactgagcgagattggtccac 122
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 109 TTCGCCAATTTGGCTTCGAGGCTCAGCGCATGCTCCACTTCGGCGGATTGGCTCCAC 168
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 123 atggacatcatgtagggcatttgcctcccaatttaactatgagctccagttatgaa 182
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 169 ATGGACATCATGATGGCATTTTGTCCCAATTTAACTATTGGCGCTCCAGTTATTGAA 228
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 183 agttgagaagacacacaaagcatattgtagtgcactatggttacaacacctt 242
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 229 AGTTGAGAGAAAGCACAAAGGATTTGGATGTCTCCTTATGTTACAAATCCTCTT 288
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 243 gatctgttgaaccttggcacaagctggtctctggtttacattcagtagagaca 302
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 289 GATTATGTTGAGCCCTTGGCAAAAGCTGCTCTCTGTTTACATTTCAATGTAAGACA 348
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 303 tcaaaagaactgaagaactatccaagaatcagaatcagatgcatgattccgtg 362
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 349 TCAAAAGATTAAGTGGAGAAAGACTTATCCAAAGATTCAGATGCAATGACTCTGGT 408
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 363 gtagcataaagcctggagcccgct 388
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 409 GTACGATTAAAGCTGGAGACCCCAT 434
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

```

RESULT 2
LOCUS      A1437973              418 bp      mRNA      linear      EST 28-NOV-2001
DEFINITION      sa34c12.y1 Gm-cl004 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
                  Gm-cl004-1199 5' similar to TR:014105 014105 RIBULOSE-PHOSPHATE
                  3-EPIMERASE. ; mRNA sequence.
ACCESSION      A1437973
VERSION        A1437973.1
KEYWORDS       EST.
SOURCE         soybean.
ORGANISM       Glycine max

```

```

REFERENCE      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS        Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                  Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
                  Glycine
                  1 (bases 1 to 418)
                  Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
                  A., Bolla,B., Marra,M., Hillier,L., Kuobu,T., Martin,J., Beck,C.,
                  Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

```

```

TITLE
JOURNAL
COMMENT

```

```

Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shln,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: coultresgen.com
Insert Length: 945 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 414
POLYA=NO.

```

```

FEATURES
source

```

```

1. 418
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl004-1199"
/clone_lib="Gm-cl004"
/issue_type="root"
/lab_host="XhoI-Gold"
/notes="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:
XhoI; Root cDNA. The mRNA was isolated from entire roots
of 8 day old 'Williams' seedlings which were propagated on
paper towels with distilled water. Stratiene's cDNA
Synthesis Kit (catalog #200401) was used to synthesize the
cDNA. First-strand synthesis was performed with 5-methyl
dCTP, hence the ligated cDNA is hemimethylated.
Stratiene's first-strand synthesis primer was used
[GAGAGAGAGAGAGAGAGACTGTCGAG(T)-18]. After
second-strand synthesis, the cDNA ends were 'polyed'
with clone pfu DNA polymerase, ligated to EcoRI adapters,
and phosphorylated. The XhoI site within the first-strand
all XhoI sites in the cDNA would be protected by their
hemimethylated status. The cDNA constructs were
size-fractionated with a 500bp cutoff, using GibcoBRL Life
Technologies' cDNA Size Fractionation column. The column
eluent was then ligated into Stratiene's pBluescript II
XR Predigested vector (pBluescript II SK(+)) that had been
digested with EcoRI and XhoI, and phosphorylated). Both
the white and blue colonies appear to contain recombinant
plasmids with cDNA inserts. Blue colonies (9n-15) have been
sequenced, and possess putative cDNA inserts. This library
was constructed by Dr. Paul Kelm & Virginia H. Coryell,
Department of Biology, Box5640, Northern Arizona
University, Flagstaff, AZ 86011, Phone: 520-523-1078 (Dr.
Paul Kelm), 520-523-1372 (Virginia H. Coryell), Fax:
520-523-7500, email: paul.kelma@uau.edu,
virginia.coryell@uau.edu"

```

```

BASE COUNT      111 a      93 c      101 g      113 t
ORIGIN

```

```

Query Match      94.2%; Score 365.4; DB 9; Length 418;
Best Local Similarity 97.1%; Pred. No. 7.0e-95;
Matches 372; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```

```

OY 6 aagaagaagaagaatggaatgacacccgaaataagctccttgatgctcttcgcgactc 65
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 4 AGAAGAGAGAGAGATGGAGTGAACCGAAATTTGCTTCATGCTCTTCGACACTC 63
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 66 gccaatgtgcttcgaggtcagcgatgctccacttggcgcgagattggtccacatg 125
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 64 GCCAATTTGGCTTCGAGGCTCAGCGCATGCTCCACTTTCGGCCCGATTGGCTCCACATG 123
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 126 gacatcatgtagggcatttgcctcccaatttaactatggtgctccagttatgaagt 185
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```



```

|||||
Db 124 GACATCATGATGGCATTTTGTCCCAATTATTAACATTTAGCCCTCCAGTATTGAAAGT 183
Qy 186 ttgaagaacacacaaaggacatatttgatctgtcaccttatggttacaatcctcttgat 245
Db 184 TTGAGAAGACACAAAGGATATTGATGTCACCTTATGTTACAAATCCTCTTGAT 243
Qy 246 tatgttgaaaccttgacaaagctggtctctggttttaacattcaacgtagagacatca 305
Db 244 TATGTTGAGCCCTTGGCAAAAGCTGTGCTTTCGTTTACATTCATGATGAGACATCA 303
Qy 306 aaagataactggaagaactatccaaagaatcacatgcacatgcatctgctgtgta 365
Db 304 AAAGATTAAGTGGAGAACTTATTCACAAAGATCAAGTCACATGACATGCTCTGGTGA 363
Qy 366 gcatlaaagccttggaacccct 388
Db 364 GCATTAAGCCTTGGAACCCCAT 386

```

```

RESULT 3
BM309659 575 bp mRNA linear EST 02-JAN-2002
LOCUS sak55b08.y1 Gm-cl036 Glycine max cDNA clone SOYBEAN CLONE ID:
DEFINITION Gm-cl036-7767 5' similar to TR:Q9SE42 Q9SE42 D-RIBULOSE-3-PHOSPHATE
3-EPIMERASE ; mRNA sequence.
ACCESSION BM309659
VERSION BM309659.1 GI:18041365
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max

```

```

REFERENCE 1 (bases 1 to 575)
AUTHORS Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna
,A., Bolla,B., Matra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
JOURNAL Contact: Shoemaker R/Public Soybean EST Project
COMMENT Public Soybean EST Project
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watscn.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccuteresgen.com web site:
www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 418.
Location/Qualifiers
1. 575
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl036-7767"
/clone_1lb="Gm-cl036"
/tissue_type="somatic embryos cultured on MSD 20"
/lab_host="DH10B"
/note="Vector: pSPORT1. Site 1: NotI. Site 2: SalI. This
cDNA library was constructed from mRNA isolated from
somatic embryos (age ranging from 2 months to 9 months)
cultured on MSD 20. The library was prepared using the
Life Technologies pSuperScript cDNA library construction
kit. Complementary DNA was synthesized from mRNA using a
poly (dT) sequence with a NotI restrictions site. SalI
linkers adapters were ligated to the blunt-ended cDNA

```

FEATURES

source

fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E. coli Electromax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu

Query Match 90.9%; Score 352.8; DB 10; Length 575; Best Local Similarity 98.1%; Pred. No. 3.8e-91; Matches 357; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

Qy 25 ttgaccggaataagctccttcgcatgctctcttcgacatcgcgaattggtcttcgag 84
Db 1 TGACACCGAATAATGCTCCTTCATGCTCTCTCCAGCTTCGCAATTTGGCTTCGAGG 60
Qy 85 ctacagcgcattcctcacccttcgcgcgcatggtctcacaatgacatcatgattgagcatt 144
Db 61 CTCAGCGCATGCTCCACTTCGCGCGATGCTCCACATGACATCATGATGATGGCAT 120
Qy 145 ttgtcccaatttaactatctggtcccgcttatctgaagtttgagaagacacacaagg 204
Db 121 TTGTCCCAATTTAACTATTGGCGCTCCAGTTATGAAAGTTGAAACACACAAAGG 180
Qy 205 cataatgtagtgcacacctatggttaacaatcctcttgattatgttgaaccttgcaa 264
Db 181 GATATTGGATTTGCTACACTTATGTTACAAATCTCTTGATTATGTTGACCCCTTGCAA 240
Qy 265 aagctggtctcttggtttacatttcacgttagagacatcaaaagataactgtgaagac 324
Db 241 AACCTGCTGCTCTGCTTACATTTACATGACATCAAAAGATTAACCTGGGAAGAC 300
Qy 325 ttatccaagaatcaagtcacatgcacatgcatgattctctggttagatcaagcctggacc 384
Db 301 TTTATCCAAAGATCAAGTCACATGATGATGCTCTGCTGTAGCATTAAGCCCTGGACCC 360
Qy 385 ccgt 388
Db 361 CCAT 364

```

```

RESULT 4
BG316079 374 bp mRNA linear EST 28-NOV-2001
LOCUS sab55g09.y1 Gm-cl043 Glycine max cDNA clone GENOME SYSTEMS CLONE
DEFINITION ID: Gm-cl043-4170 5' similar to TR:Q9SE42 Q9SE42
D-RIBULOSE-5-PHOSPHATE 3-EPIMERASE ; mRNA sequence.
ACCESSION BG316079
VERSION BG316079.1 GI:13125509
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max

```

```

REFERENCE 1 (bases 1 to 374)
AUTHORS Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Corryell,V., Khanna
,A., Bolla,B., Matra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
JOURNAL Contact: Shoemaker R/Public Soybean EST Project
COMMENT Public Soybean EST Project
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

```


Fax: 314 286 1810
 Email: estewatson.wustl.edu
 This clone is available through: Resgen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-553-4363 or contact via email: coudresgen.com
 High quality sequence stop: 351.

FEATURES

source

1..374
 Location/Qualifiers
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl043-4170"
 /clone_1ib="Gm-cl043"
 /tissue_type="hypocotyl and plumule, germinating seeds"
 /lab_host="DH10B"
 /note="Vector: p7T73pac (Pharmacia); Site_1: EcoRI;
 Site_2: NotI; This cDNA library was constructed from mRNA
 isolated from hypocotyl and plumule tissues of seeds
 germinated for three days of the cultivar Williams.
 Complementary DNA was synthesized from mRNA using a primer
 consisting of a poly(dT) sequence with a NotI restriction
 site. EcoRI adapters were ligated to the blunt-ended cDNA
 fragments followed by digestion with EcoRI and NotI. The
 EcoRI-NotI restriction site of the p7T73-pac vector. The
 ligated cDNA fragments were transformed into DH10B host
 cells (Gibco BRL). This library was constructed by Dr.
 Randy Shoemaker."

BASE COUNT 104 a 80 c 89 g 101 t
 ORIGIN

Query Match

Best Local Similarity 87.2%; Score 338.2; DB 10; Length 374;
 Matches 343; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 3 agaagaagaagaatggaatgacacggaagaatagctccttcgatgctcttcgcgac 62
 Db 24 AGAAGAAGAAGAGATGGAGTGCACCGAAATTCCTCTCGATGCTCTTCCGAC 83
 Qy 63 ttcgcaattggtctcgaagctcgaagcagcagctgctcgaagcagcagcttcac 122
 Db 84 TTCGCCAATTGGCTTCGAGGCTCAGCGCATGCTCCACTTCGGCGCGATGGCTCCAC 143
 Qy 123 atgacatcatgcatggaatcttgcctcccaatttaactatgagcgtccagttatgaa 182
 Db 144 ATGACATCATGATGGGCAATTTGTCGCCAATTTAACTATTGGCGCTCCAGTTATGAA 203
 Qy 183 agttggaagaagcacaaagcatttgatgtcaactatggttacaactcctct 242
 Db 204 AGTTGAGAAAGCACAAAGGATATTGGATTGTACCTTATGTTACAAATCCCTCTT 263
 Qy 243 gataatgtgaacctgtgcgaagaagctgtgctctgtttacatttcagctgaagaga 302
 Db 264 GATTATGTGTAGCCCTTGGCAAAAGCTGCGTCTTCGTGTTTCAATTTCAATGAGAGACA 323
 Qy 303 tcaaaagataactcggaaagaactatccaaagaatcaagtcacatgcatg 353
 Db 324 TCAAAAGATAACTGGGAAAGACTTATCCAAAGATCAAGTCATGCGCATG 374

RESULT 5

AV412302 403 bp mRNA linear EST 23-MAY-2000
 LOCUS AV412302 Lotus japonicus young plants (two-week old) Lotus
 DEFINITION japonicus cDNA clone MM217c10_r 5', mRNA sequence.
 ACCESSION AV412302
 VERSION AV412302.1 GI:7741466
 KEYWORDS EST.

SOURCE

Lotus japonicus.

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;

REFERENCE

1 (bases 1 to 403)
 AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
 TITLE Generation of 7137 non-redundant expressed sequence tags from a
 legume, Lotus japonicus
 JOURNAL DNA Res. 7 (2), 127-130 (2000)
 MEDLINE 20277479

COMMENT

Contact: Yasukazu Nakamura
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: ynakam@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

source

1..403
 Location/Qualifiers
 /organism="Lotus japonicus"
 /db_xref="taxon:34305"
 /clone="MM217c10_r"
 /clone_1ib="Lotus japonicus young plants (two-week old)"
 /dev_stage="young plants (two-week old)"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI; Isolate=MIYAKOJIMA MG-20"

BASE COUNT 113 a 93 c 91 g 106 t
 ORIGIN

Query Match

Best Local Similarity 82.7%; Score 320.8; DB 9; Length 403;
 Matches 340; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 4 gaaagaagaagaatggaatgacacggaagaatagctccttcgatgctcttcgcgact 63
 Db 32 GTAACAACAAAGAGATGGAGTGCACCCAAATTCCTCGATGCTTCATATGGACT 91
 Qy 64 tgcgaattggtctcgaagcgtccagcgcatgctcgaacttcggcgcgatggtccaca 123
 Db 92 TCCCAATTTGGCTTCGAGGCTCAACGCATCTGATTATAGCCCGATGGCTCCACA 151
 Qy 124 tggacatcatgcatggtgcaattgtccccaatttaactatgagcgtccagttatgaa 183
 Db 152 TGGACATCATGATGGGCACTTGTCCCAATTAACTATTTAGCTCGCTCATTTGAAA 211
 Qy 184 gtttgaagaagacacaaagcagcatttgatgtcaccattatggttacaactccttg 243
 Db 212 GTTTGGAAGAAGACAAAGGCAATCTGATGTCACCTTATGTTCAAAATCCTCTTG 271
 Qy 244 attatgtgaacctgtgcgaagaagctgtgctctgtgttaccttaccgttaagagacat 303
 Db 272 ATTATGTTAGCCTTTGGGAAGACAGCTGCTTCGTTTACATTCATGTAGAGGCAT 331
 Qy 304 caaagaataactcggaaagaactatccaaagaatcaagtcacatgcatgcatgctgtg 363
 Db 332 CAAAGATTAACGTGAAAGAGATTATCCAAAGATTAAAGTACAGGCGCATAGGCTGTG 391
 Qy 364 tagcattaagc 375
 Db 392 TAGCATTTAAAGC 403

RESULT 6

AI442690 405 bp mRNA linear EST 28-NOV-2001
 LOCUS AI442690
 DEFINITION sa85b04.y1 Gm-cl004 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-cl004-6056 5' similar to TR:014105 O14105 RIBULOSE-PHOSPHATE
 3-EPIMERASE.; mRNA sequence.
 ACCESSION AI442690
 VERSION AI442690.1 GI:4297942
 KEYWORDS EST.

SOURCE

soybean.

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL04R cells."

BASE COUNT

183 a 110 c 131 g 150 t

ORIGIN

Query Match 79.3%; Score 307.8; DB 10; Length 574;
Best Local Similarity 87.7%; Pred. No. 3.7e-78;
Matches 336; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 6 aagaagaagaatgggaatgacacggaataatgctcctcgtcgtcctcgcagctc 65
D 32 AAGAAGAAGAATGGGAGTGCACGCAAAATAGCTCTTCATGCTATCATCATATTT 91
QY 66 gccaatctgcttcgcagcagcagcagcagcagcagcagcagcagcagcagcagc 125
D 92 GCTAATTTGGCTTCGACAGCTCATGATGATCATTTACGGCGGATGATGCTTCACATG 151
QY 126 gacatcatgagtgagcttctcccaattactatgagcagcagcagcagcagcagc 185
D 152 GATATCATGATGGCATTTTGTCTTAACTTAATGAGCGCTCCATCATTTGAAGT 211
QY 186 ttgagaagaacacacaaagacatattgattgacacattgattgacacattgattg 245
D 212 TTGAGAAACACACAGAGGATATCTGACCTGATGATGATGATGATGATGATGATG 271
QY 246 tatgtgaaccttggaagacgctgctcgtctcgtctcgtctcgtctcgtctcgtctc 305
D 272 TATGTTGAACCTTTGGCAAAAGCTGTGCTCTGTTTACATTTTACATTTTACATTT 331
QY 306 aagaataactggaagaactatccaagaatcaagacatgacatgacatgacatgac 365
D 332 AAGACAACTGGAAAGAACTTATCCAAATTTAAGTACACGCGCATGAGCGCTGCTGTA 391
QY 366 gcattaagccttggagcccgct 388
D 392 TCGATTAAGCCTGGAACATCCGT 414

RESULT 8

LOCUS

AM682881

641 bp mRNA

linear

EST 15-JUN-2000

DEFINITION

MF001C10LF1081 Developing leaf Medicago truncatula cDNA clone

ACCESSION

AM682881

641 bp mRNA

linear

EST 15-JUN-2000

VERSION

AM682881.1

GI:7557597

KEYWORDS

EST

SOURCE

barrel medic.

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

ORGANISM

barrel medic.

Medicago truncatula

REFERENCE

1 (bases 1 to 641)

AUTHORS

Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Imman, J.T., Weller, J.W., and May, G.D.

TITLE

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

JOURNAL

Medicago truncatula leaf library

COMMENT

Unpublished (2000)

CONTACT

May CD

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7391

Fax: 580 221 7380

Email: gdmay@noble.org

Insert Length: 641

Plate: 001 row: C column: 10

Seq primer: TCACACAGGAAACACCTATGAC.

Location/Qualifiers

1..641

/organism="Medicago truncatula"

FEATURES

source

/organism="Medicago truncatula"

/db_xref="taxon:3880"
/clone="MF001C10LF"
/clone_1lb="Developing leaf"
/tissue_type="leaf"
/dev_stage="Pooled developmental"
/note="Vector: Lambda Zap; Contains a mixture of very young, developing, mature and senescing leaves."

BASE COUNT 201 a 130 c 139 g 171 t

ORIGIN

Query Match 79.3%; Score 307.8; DB 9; Length 641;
Best Local Similarity 87.7%; Pred. No. 3.9e-78;
Matches 336; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 6 aagaagaagaatgggaatgacacggaataatgctcctcgtcgtcctcgcagctc 65
D 126 AAGAAGAAGAATGGGAGTGCACGCAAAATAGCTCTTCATGCTATCATCATATTT 185
QY 66 gccaatctgcttcgcagcagcagcagcagcagcagcagcagcagcagcagcagc 125
D 186 GCTAATTTGGCTTCGACAGCTCATGATGATGATGATGATGATGATGATGATGATG 245
QY 126 gacatcatgagtgagcttctcccaattactatgagcagcagcagcagcagcagc 185
D 246 GATATCATGATGGCATTTTGTCTTAACTTAATGAGCGCTCCATCATTTGAAGT 305
QY 186 ttgagaagaacacacaaagacatattgattgacacattgattgacacattgattg 245
D 306 TTGAGAAACACACAGAGGATATCTGACCTGATGATGATGATGATGATGATGATG 365
QY 246 tatgtgaaccttggaagacgctgctcgtctcgtctcgtctcgtctcgtctcgtctc 305
D 366 TATGTTGAACCTTTGGCAAAAGCTGTGCTCTGTTTACATTTTACATTTTACATTT 425
QY 306 aagaataactggaagaactatccaagaatcaagacatgacatgacatgacatgac 365
D 426 AAGACAACTGGAAAGAACTTATCCAAATTTAAGTACACGCGCATGAGCGCTGCTGTA 485
QY 366 gcattaagccttggagcccgct 388
D 486 TCGATTAAGCCTGGAACATCCGT 508

RESULT 9

LOCUS

BG455207

654 bp mRNA

linear

EST 19-MAR-2001

DEFINITION

MF102610PL1F1082 Phosphate starved leaf Medicago truncatula cDNA

ACCESSION

BG455207

654 bp mRNA

linear

EST 19-MAR-2001

VERSION

BG455207.1

GI:13378532

KEYWORDS

EST

SOURCE

barrel medic.

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

ORGANISM

barrel medic.

Medicago truncatula

REFERENCE

1 (bases 1 to 654)

AUTHORS

Liu, J., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Imman, J.T., Weller, J.W., May, G.D., and Harrison, M.J.

TITLE

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

JOURNAL

Medicago truncatula phosphate-starved leaf library

COMMENT

Unpublished (2000)

CONTACT

Harrison MJ

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7325

Fax: 580 221 7380

Email: mjharrison@noble.org

Insert Length: 654

Std Error: 0.00

Plate: 102 row: G column: 10

Seq primer: TCACACAGGAACACGCTATGAC

FEATURES	Location/Qualifiers
source	1. .654

/organism="Medicago truncatula"
 /db_xref="taxon:3880"
 /clone="NF102610PL"
 /clone_id="Phosphate starved leaf"
 /tissue_type="leaf"
 /dev_stage="trifoliolate"
 /note="Vector: lambda Zap; At the trifoliolate stage, M.
 truncatula plants were transplanted to phosphate-free sand
 and grown for a further 30 days. During this 30 day
 period, the plants were fertilized twice weekly with 1/2
 Hoaglands solution containing only 20um potassium
 phosphate. RNA was prepared from above ground tissues."

BASE COUNT	202 a	136 c	143 g	172 t	1 others
ORIGIN					

Query Match	79.3%	Score 307.8	DB 10	Length 654
Best Local Similarity	87.7%	Pred. No. 3.9e-78		
Matches 336	Conservative 0	Mismatches 47	Indels 0	Gaps 0

QY	b	aaggaagaagaagaatgaggggaatgaacacggaataatagctcttcgaatgccttcacgacttc	65
Db	117	AAAAGAGAAAGGAGGGAGGTGACACCAGAAATAGCTCCTTGAGATGATCATCAGATT	176
QY	6	gcceatttgccttcgagagctcagcgcatgctccaattggcgcgcatgtgtccacatg	123
Db	177	GCTATATTGGCTTCGGAAGCTCATCGTATCATTAATACGGCGCTGATTTGGCTTCACATG	236
QY	126	gacatactgataaggcaatttgcgcccaatttaactatttggcgctccagttatgtgaagt	185
Db	237	GATATCATGGATGGGCGATTTGTGCCATTAATTAACCTTGGGCGCTCCAAATCATTTGAAGT	296
QY	186	tttgaagaagcacacaagaagcatttggatgtgcaccttattggttacaacatcccttgat	245
Db	297	TTTGGAAGAAGCACACAGAGGCGATATCTGAGCTGTACCTGTATGGTTACAAATTCACCTTGAT	356
QY	246	tatgttgaacctcttgacaaagaagctggtgcttcgtgttttaacttccagtagaagacata	305
Db	357	TATCTTGACCTTTGGGAAAAAGCTGGTGGCTTCCTGGTTTACATTTCATATAGAGACATCA	416
QY	306	aaagataacttggaaagaactatccaaagaatcaagttacatgcatgcatgtatcctggtgta	365
Db	417	AAAGACACACTGGAAAGAACTTATCCAAAATATTAAAGTCACACGGCATAGAGCCGTGCTGA	476
QY	366	gcattaagcttggagaccccgct	388
Db	477	TCGATTAAGCCTGGAGACATCCGT	499

RESULT	10
LOCUS	BF645593
DEFINITION	683 bp mRNA linear EST 20-DEC-2000 NF016H02C1P10Z7 Elicited cell culture Medicago truncatula CDNA
ACCESSION	NF016H02EC 5', mRNA sequence.
VERSION	BF645593
KEYWORDS	BF645593.1 GI:11910722
SOURCE	EST.
ORGANISM	barrel medic. <i>Medicago truncatula</i> Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosoids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; <i>Medicago</i> . 1 (bases 1 to 683) Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores,H.R., Imman,J.T., Weller,J.W. and May,G.D. Expressed Sequence Tags from the Samuel Roberts Noble Foundation - Center for Medicago Genomics Research Unpublished (2000) Contact: Dixon RA
JOURNAL	
COMMENT	

Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: rdixon@noble.org
Insert length: 683 std Error: 0.00
plate: 016 Row: H Column: 02
Seq primer: TCACACAGGAAACACCTATGAC.

FEATURES
source

1. .683

```

/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF016H02EC"
/clone_lib="Elicited cell culture"
/tissue_type="Cell cultures derived from root tissues"
/dev_stage="Cell suspensions were subcultured every 14
days. Cells were induced six days after subculture"
/notes="vector: lambda zap; Cells were induced with yeast
cell wall extracts equivalent to 50ug/ml glucose in the
final concentration. Samples were taken at 0.5, 1, 12 and
24 hours after induction. Equal amounts of RNA from each
time point were pooled and used for mRNA isolation."

```

BASE COUNT	209 a	138 c	153 g	183 t
ORIGIN				

Query Match	79.3%	Score 307.8	DB 10	Length 683
Best Local Similarly	87.7%	Pred. No. 4e-78		
Matches 336	Conservative 0	Mismatches 47	Indels 0	Gaps 0

QY	6	aagaaagaagaagatgaggaaatgaaacccaagaatagctcttcgatgctcttcacgactc	65
Db	125	AAAAGAGAGAAAGATGGGAGTACACCCGAAAAThAGCTCTCTTCGATGCTATATCATCAGATTTT	184
QY	66	gccaatctgctcttcgaaggtcgaacgaatgctccacttcggcgcgaattgctccacatg	125
Db	185	GCTATTTGGCTTCCGAAGCTCATCTGATATGATCAATTACGGCGCTGATTTGGCTTCACATG	244
QY	126	gaactcaatgagtagggcaattttgtccccaatttaactctttggcgtctccagttattgaagt	185
Db	245	GATATCATGATGGGCGCAATTGTGCTCAATTTAACTATGGGCGCTCCAAATCATTTGAAGT	304
QY	186	tttgaagaagcacacaagaagcacaatttggatgtcacttaagtgtacaactcctctgac	245
Db	305	TTGAGAGAAAGCACAAGAGGCTATCTTGACGTCTACCTGATGGTTACAAATTCACCTTGAT	364
QY	246	tatgttgaacctcttgcaaaagctgtgtctcttgatttcaattccagtagagacatca	305
Db	365	TATGTGAACTTTGGGAAAAGCTGTGGTCTCTGTTTACATTTCAATATAGAGACATCA	424
QY	306	aaagataacttgaaagaactatcccaagaatcaagtcacatgycatgattctctgtgtta	365
Db	425	AAAAGACAACCTGGAAAGAACTTATCCAAAATATTAACTACACGGCATGAGGCCCTGGGTGA	484
QY	366	gcattaaagcttgagaccocgct	388
Db	485	TCGATAAAGCCTGGACATCTCGT	507

RESULT 11					
LOCUS	BC451782				
DEFINITION	BC451782	597 bp	mRNA	linear	EST 16-MAR-2001
ACCESSION	NF094040	PT1031	Drought	Medicago	truncatula clone NF094040.1
VERSION	5'				5', mRNA sequence.
KEYWORDS	BC451782				
SOURCE	BC451782.1	GI:13370576			
ORGANISM	EST.				
	barrel medic.				
	Medicago truncatula				
	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;				
	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:				
	Rosidae; eustosids I; Fabales: Fabaceae; Papilionoideae; Trifoliaceae;				

REFERENCE 1 (bases 1 to 597)
 AUTHORS Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
 TITLE Flores,H.R., Iman,J.T., Weller,J.W. and May,G.D.
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula drought library
 JOURNAL Unpublished (2000)
 COMMENT Contact: May GD
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 221 7391
 Fax: 580 221 7380
 Email: gdmay@noble.org
 Insert Length: 597 Std Error: 0.00
 Plate: 094 row: D column: 04
 Seq primer: TCACACGGAACAGCTATGAC.

FEATURES
 source
 1..597
 /organism="Medicago truncatula"
 /db_xref="taxon:3880"
 /clone="NF094D04DT"
 /clone.lib="Drought"
 /tissue.type="Plantlets"
 /dev_stage="Pooled timepoints"
 /note="Vector: Lambda Zap; Contains a mixture of entire
 plantlets harvested in a series of days-post-watering
 timepoints."

BASE COUNT 188 a 115 c 135 g 154 t 5 others
 ORIGIN

Query Match 76.5%; Score 296.8; DB 10; Length 597;
 Best Local Similarity 87.5%; Pred. No. 5,7e-75;
 Matches 336; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 6 aagaagaagaatggaatgacac-gaaatagctcttcgctcctccgactt 64
 DB 42 AAAAGAAAGATGGAGTGAACCTGAAATAGCTCTTGATGCTATCAGATT 101
 QY 65 ccccaattggtcttcgaggtcagcgatgctcacttcggcgcgattggtccacat 124
 DB 102 TGTATTTTGGCTCCGAAAGCTCATCGATGATCAATTACGGCGCTGATGGCTTCACAT 161
 QY 125 ggaacatcagatggtgcatcttgcctcccaattactatcttggtcgtccagtatgaag 184
 DB 162 GGATATCATGATGGGCAATTTGTCCCTAATTAACTATGGCGCTCCAAATCATTTGAAAG 221
 QY 185 ttgggaagacacaaaggcatattggtatgtcaccttatggtttacaatcctctga 244
 DB 222 TTGGAAGAACACAGAGGCTATCTGACTGTCACCTGATGATGTTCAAAATCCACTTGA 281
 QY 245 tatgtgaacccttgcaaaagctggtcttcgttttcaattcaatcgttaagaacatc 304
 DB 282 TTATGTGACCTTTGGAAAAAGCTGTGCTTCTTGATTACATTATATATAGAGAAATC 341
 QY 305 aaagataactcggaaagaactatccaaagaatcaagtcacatgcatgcatctcgtgt 364
 DB 342 AAAAGACAACTGGAAAGAACTTATCCAAAATATTAAGTACACGCGATGAGGCTGTGT 401
 QY 365 agcatlaaagccttggaacccctt 388
 DB 402 ATCGATTAAGCCTGGAAACATCCGT 425

RESULT 12
 BE325053 613 bp mRNA linear EST 21-DEC-2000
 LOCUS NF119G01ST1P1006 Developing stem Medicago truncatula cDNA clone
 DEFINITION NF119G01ST 5', mRNA sequence.
 ACCESSION BE325053
 VERSION BE325053.2 GI:11935879
 KEYWORDS EST.

SOURCE
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 REFERENCE 1 (bases 1 to 613)
 AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
 ,C.J., Flores,H.R., Iman,J.T., Weller,J.W., May,G.D. and Dixon
 ,R.A.
 TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula stem library
 JOURNAL Unpublished (2000)
 COMMENT On Jul 14, 2000 this sequence version replaced gi:9198914.
 Contact: Dixon RA
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 221 7302
 Fax: 580 221 7380
 Email: radixon@noble.org
 Medicago Genome Initiative accession: MGI:S.15815
 Insert length: 633 Std Error: 0.00
 Plate: 119 row: G column: 01
 Seq primer: TCACACGGAACAGCTATGAC.
 Location/Qualifiers
 1..613
 /organism="Medicago truncatula"
 /db_xref="taxon:3880"
 /clone="NF119G01ST"
 /clone.lib="Developing stem"
 /tissue.type="stem"
 /dev_stage="Pooled developmental"
 /note="Vector: Lambda Zap; Contains a mixture of
 internodal stem segments"

BASE COUNT 186 a 125 c 134 g 168 t
 ORIGIN

Query Match 76.5%; Score 296.8; DB 10; Length 613;
 Best Local Similarity 87.5%; Pred. No. 5,8e-75;
 Matches 336; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 6 aagaagaagaatggaatgacacgaaatagctccttcgctcctccgactt 65
 DB 104 AAAAGAAAGATGGAGTGAACGAAATAGCTCTTCATGCTATCAGATT 163
 QY 66 gccaattggtcttcgaggtcagcgatgctcacttcggcgcgattggtccacatg 125
 DB 164 GCTAATTTGGCTCCGAAAGCTCATCGTATGATCAATTACGGCGCTGATGGCTTCACATG 223
 QY 126 gacatcagatggtgcatcttgcctcccaattactatcttggtcgtccagtatgaagtc 185
 DB 224 GATATCATGATGGGCAATTTGTCCCTAATTAACTATGCGCGCTCCAAATCATTTGAAAGT 283
 QY 186 ttgagaagaacacaaaggcatattggtatgtcaccttatggtttacaatcctctgat 245
 DB 284 TTGGAAGAACACAGAGGCTATCTGACTGTCACCTGATGATGTTCAAAATCCACTTGAAT 343
 QY 246 tatgtgaacccttgcaaaagctggtcttcgttttcaattcaatcgttaagaacatca 305
 DB 344 TATGTGAACTTTTGGAAAAAGCTGTGCTTCTTGATTACATTATATATAGAGAAATCA 403
 QY 306 aaagataactcggaaagaactatcc-aaagaatcaagtcacatgcatgcatctcgtgt 364
 DB 404 AAAGACAACTGGAAAGAACTTATCCAAAATATTAAGTACACGCGATGAGGCTGTGT 463
 QY 365 agcatlaaagccttggaacccctt 388
 DB 464 ATCGATTAAGCCTGGAAACATCCGT 487

RESULT 13

LOCUS	BF634692	666 bp	mRNA	linear	EST 19-DEC-2000
DEFINITION	BF634692		NF064B05D1F1044 Drought Medicago truncatula cDNA clone NF064B05DT		
ACCESSION	BF634692				
VERSION	BF634692.1				
KEYWORDS	EST.				
SOURCE	barrel medic.				
ORGANISM	Medicago truncatula				
REFERENCE	1 (bases 1 to 666)				
AUTHORS	Torrez-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Imman, J.T., Weller, J.W. and May, G.D.				
TITLE	Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula drought library				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: May GD Plant Biology Division The Samuel Roberts Noble Foundation 2510 Sam Noble Parkway, Ardmore, OK 73402, USA Tel: 580 221 7391 Fax: 580 221 7380 Email: gdmay@noble.org Insert Length: 666 Std Error: 0.00 Plate: 064 row: B column: 05 Seq primer: TCACACAGAAACAGCTATGAC. Location/Qualifiers 1..666 /organism="Medicago truncatula" /db_xref="taxon:3880" /clone_lib="NF064B05DT" /clone_lib="Drought" /lissue_type="Plantlets" /dev_stage="Pooled timepoints" /note="Vector: lambda Zap, Contains a mixture of entire plantlets harvested in a series of days-post-watering timepoints."				
FEATURES	source				
BASE COUNT	189 a 127 c 165 g 182 t 3 others				
ORIGIN					
Query Match	75.6%; Score 293.2; DB 10; Length 666;				
Best Local Similarity	87.8%; Pred. No. 6.5e-74;				
Matches 330; Conservative 0; Mismatches 45; Indels 1; Gaps 1.					
14	aaagttgggaatgaacaccggaataatgcttccttcgtagtctcttcgcgaattgccaatt 73				
2	AAAGATGGGAGATNNACCGAAATAGCTCTTCGATGCTATCATCAGATTTGCTAATT 61				
74	gagcttcgagagctcagcgatcgcacttc-ggcgccgattggtccacatgagacata 132				
62	GGCTTCGGAAGCTCATCGATGATCAATATTACGGCGCTGATTTGGCTTCACATGATATCA 121				
133	tggaatgggcatcttctccccaatttaactatctgagcgctcgcattatgaaattgagaa 192				
122	TGGATGGGCGATTGTCCTTATTTAACTATGGGGCGCTCCATCAFTGAAAGTTTGAGAA 181				
193	agcaacaaaggcatatttgattgtcacctatgagttacaaatccctctgaltatgttg 252				
182	AGCACACGAGGCAATATCGACCTGCACCTCATGATGTTCAAAATCCACTTGATTATGTTGG 241				
253	aacctctggcaaaagctgtgctcttgattcaactttacgtaagaaagataaagata 312				
242	AACCTTTGGGAAAACCTGTGCTTCTTGTTTACATTTATATATGAGACATCAAAAAGACA 301				
313	actggaagaacttatccaagaagatcaagatcaatgacatgacatgattcctgtgtagactaa 372				
302	ACTGGAAGAAGACTTATTCACAAATATTAAGTACACAGGCGATGAGGCGTGTATGATATA 361				
373	agcctggagaccccgct 388				

DB	362	AGCCTGGAACATCCCT	377
RESULT 14			
LOCUS	AM348508/c		
DEFINITION	AM348508	736 bp	mRNA linear EST 04-OCT-2000
ACCESSION	GM10002B12C3	Gm-r1021	Glycine max CDNA clone Gm-r1021-509 3', mRNA
VERSION			
KEYWORDS	sequence.		
SOURCE	AM348508.1	GI:6846218	
ORGANISM	soybean.		
REFERENCE	Glycine max		
AUTHORS	Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.		
TITLE	1 (bases 1 to 736)		
JOURNAL	Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V., Erpelting, J., Rapp, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H. A Functional Genomics Program for Soybean (NSF 9872365)		
COMMENT	Unpublished (1999) other ESTs: A1437973 Contact: Vodkin, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872365) Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics University of Illinois Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA Tel: (217) 244-6147 Fax: (217) 333-4582 Email: l-vodkin@uiuc.edu This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134. For further information call: (800) 430-0030 or (314) 427-3324 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genome systems.com web site: www.genomesystems.com Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3', Location/Qualifiers 1..736 /organism="Glycine max" /cultivar="Williams" /db_xref="taxon:3847" /clone="Gm-r1021-509" /clone_id="Gm-r1021" /issue_type="root" /lab_host="XLI0-Gold" /note="Vector: plusescript II XR, Site1: EcoRI, Site2: XhoI; Library Gm-r1021 is a sequence-driven, rereaked set of the original library Gm-cl004 which was prepared from root cDNA. The mRNA was isolated from entire roots of 8 day old 'Williams' seedlings which were propagated on paper towels with distilled water. Stratagene's cDNA Synthesis Kit (catalog #200401) was used to synthesize the cDNA. The Gm-cl004 library was constructed by Dr. Paul Keim & Virginia H. Coryell, Department of Biology, Box6640, Northern Arizona University, Flagstaff, AZ 86011, email: paul.keim@uau.edu, virginia.coryell@uau.edu The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Computational Biology Centers, University of Minnesota, http://www.cbc.umn.edu/ResearchProjects/Soybean/Index.html . Rereaking was performed by Genome Systems, St. Louis, http://www.genomesystems.com, and sequencing by the Keck Center for Comparative and Functional Genomics, University of Illinois, http://www.life.uiuc.edu/biotech/keck.html."		
BASE COUNT	197 a	164 c	131 g 215 t 29 others
ORIGIN			
Query Match			

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148563.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 58.0%; Score 225.2; DB 21; Length 1034;
Best Local Similarity 75.1%; Pred. No. 1.5e-61;
Matches 281; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 15 aagatggaatgacaccgaaatagctccttgatgtctcttcgcgaacttg 74
DB 26 aagaagaatggtgctgcgcaaaataagaccgctgctgcatcgcgaactc 85
QY 75 gcttcggaagcttaagcgtatgctccacttcgcgcgcatggtccacatgacatgacatg 134
DB 86 gcgcgcgaagccaatgcgatgacgtattggcgcaactgcgttcacatgataatg 145
QY 135 gatggcatttgcctcccaacttaactattgcgctccagttatgaagttagaag 194
DB 146 gacggcatttgcctcactatctacgatgtgctccctgcacatgagatgacgaag 205
QY 195 cacacaagaagcatttgatgtgctacattatgttacaaatcctcttgattgtgaa 254
DB 206 cacacaaactatcttgcctcactatgctgacgaaccacatgattgactgca 265
QY 255 cccttgcaaaagctgtctcttgctgtttacattcacaqtaagacataaagaataac 314
DB 266 cagatggtcaaaagctgggctcttgcgtttcaccattccacgtcgaggtgcccagaataat 325
QY 315 tggaaagaacttaccaaagaatcaagtcacatgacatgattcctgtgtagcatataag 374
DB 326 tggcaacaacttgcgcagaagattaaacttaactggtgatgagaccagcggtgctcctaag 385
QY 375 cctgggaaccccgct 388
DB 386 cctgggaacacctgt 399

RESULT 2
AAC44542
ID AAC44542 standard; DNA; 1153 BP.
XX
AC AAC44542;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays DNA fragment SEQ ID NO: 43215.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW Protein identification; signal transduction pathway; metabolic;
KW pathway; promoter; termination sequence; corn; ss.
XX
OS Zea mays subsp. mays.

XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.

RESULT 3
AAC33528
ID AAC33528 standard; DNA; 910 BP.
XX
AC AAC33528;
XX
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 3381.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148665.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR	20-AUG-1999;	990S-0149722.
PR	20-AUG-1999;	990S-0149723.
PR	20-AUG-1999;	990S-0149929.
PR	23-AUG-1999;	990S-0149902.
PR	23-AUG-1999;	990S-0149930.
PR	25-AUG-1999;	990S-0150565.
PR	26-AUG-1999;	990S-0150566.
PR	27-AUG-1999;	990S-0150884.
PR	27-AUG-1999;	990S-0151065.
PR	27-AUG-1999;	990S-0151066.
PR	30-AUG-1999;	990S-0151080.
PR	31-AUG-1999;	990S-0151303.
PR	01-SEP-1999;	990S-0151438.
PR	07-SEP-1999;	990S-0151930.
PR	10-SEP-1999;	990S-0152363.
PR	13-SEP-1999;	990S-0153070.
PR	15-SEP-1999;	990S-0153758.
PR	16-SEP-1999;	990S-0154018.
PR	20-SEP-1999;	990S-0154039.
PR	22-SEP-1999;	990S-0154779.
PR	23-SEP-1999;	990S-0155139.
PR	24-SEP-1999;	990S-0155486.
PR	28-SEP-1999;	990S-0155659.
PR	29-SEP-1999;	990S-0156458.
PR	04-OCT-1999;	990S-0156599.
PR	05-OCT-1999;	990S-0157117.
PR	06-OCT-1999;	990S-0157753.
PR	07-OCT-1999;	990S-0157865.
PR	08-OCT-1999;	990S-0158029.
PR	12-OCT-1999;	990S-0158232.
PR	13-OCT-1999;	990S-0158369.
PR	13-OCT-1999;	990S-0159283.
PR	13-OCT-1999;	990S-0159284.
PR	13-OCT-1999;	990S-0159295.
PR	14-OCT-1999;	990S-0159329.
PR	14-OCT-1999;	990S-0159330.
PR	14-OCT-1999;	990S-0159331.
PR	14-OCT-1999;	990S-0159637.
PR	14-OCT-1999;	990S-0159638.
PR	18-OCT-1999;	990S-0159584.
PR	21-OCT-1999;	990S-0160741.
PR	21-OCT-1999;	990S-0160767.
PR	21-OCT-1999;	990S-0160770.
PR	21-OCT-1999;	990S-0160814.
PR	21-OCT-1999;	990S-0160815.
PR	22-OCT-1999;	990S-0160980.
PR	22-OCT-1999;	990S-0160981.
PR	22-OCT-1999;	990S-0160989.
PR	25-OCT-1999;	990S-0161404.
PR	25-OCT-1999;	990S-0161405.
PR	25-OCT-1999;	990S-0161359.
PR	26-OCT-1999;	990S-0161360.
PR	26-OCT-1999;	990S-0161361.
PR	28-OCT-1999;	990S-0161920.
PR	28-OCT-1999;	990S-0161993.
PR	29-OCT-1999;	990S-0162142.

Db 169 tggacgagcatttgccttctaacccaagaatgctgcctctgtcatclogaggtttggga 228

QY 193 agcacacaagagcattttgatgtgcacctatgtgttaacaatcctctgattatcgtg 253

Db 229 agcacacaagcgtatcctgtatgtccaactaatggttgacgaaccctagtattatg 288

QY 253 aacctgtgcaaaagctgctgtcctcgtgtttacaatttcacgtlaaagacacaaagata 312

Db 289 atcagatgtgtcaaaagctgcggcgctcgtgttcacatccaagcttgaggttgcccaagga 348

QY 313 actggaagaactatcccaagaagataaagatcacatggtcatgttuccgtgttaagcattaa 372

Db 349 atgtgcaagaactgtgaaagaagatctaaggtcgtgcgaatgaagccagttgtgctctaa 408

QY 373 agccttgagaccccgct 388

Db 409 agccttgaaacacgtgt 424

	RESULT	4
XX	ID	AAC43724
XX		AAC43724 standard; DNA; 462 BP.
XX	AC	AAC43724;
XX	DF	18-OCT-2000 (first entry)
XX	DE	Zea mays DNA fragment SEQ ID NO: 40278.
KM		Hybridisation assay; genetic mapping; gene expression control;
KW		protein identification; signal transduction pathway; metabolic;
XN		pathway; promoter; termination sequence; corn; ss.
OS		Zea mays subsp. mays.
XX	PX	EP1033405-A2.
XX	PN	
XX	PD	06-SEP-2000.
XX	PE	
XX		25-FEB-2000; 2000EP-0301439.
PR		25-FEB-1999; 99US-0121825.
PR		05-MAR-1999; 99US-0122180.
PR		09-MAR-1999; 99US-0123548.
PR		23-MAR-1999; 99US-0125788.
PR		25-MAR-1999; 99US-0126264.
PR		29-MAR-1999; 99US-0126785.
PR		01-APR-1999; 99US-0127462.
PR		06-APR-1999; 99US-0128234.
PR		08-APR-1999; 99US-0128714.
PR		16-APR-1999; 99US-0129845.
PR		19-APR-1999; 99US-0130077.
PR		21-APR-1999; 99US-0130449.
PR		23-APR-1999; 99US-0130510.
PR		23-APR-1999; 99US-0130891.
PR		28-APR-1999; 99US-0131449.
PR		30-APR-1999; 99US-0132048.
PR		30-APR-1999; 99US-0132407.
PR		04-MAY-1999; 99US-0132484.
PR		05-MAY-1999; 99US-0132485.
PR		06-MAY-1999; 99US-0132486.
PR		07-MAY-1999; 99US-0132487.
PR		11-MAY-1999; 99US-0134256.
PR		14-MAY-1999; 99US-0134218.
PR		14-MAY-1999; 99US-0134219.
PR		14-MAY-1999; 99US-0134221.
PR		18-MAY-1999; 99US-0134370.
PR		19-MAY-1999; 99US-0134768.
PR		20-MAY-1999; 99US-0135124.
PR		21-MAY-1999; 99US-0135523.
PR		24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140635.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142350.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144335.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.
PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145224.
PR 23-JUL-1999; 9905-0145228.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.

PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151308.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155119.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 18-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160989.
PR 22-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 25-OCT-1999; 9905-0161405.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.


```
RESULT 8
AAA67520
ID   AAA67520 standard; DNA; 410 BP.
XX
AC   AAA67520;
XX
DT   31-OCT-2000 (first entry)
XX
DE   Pinus radiata ribulose-phosphate-3-epimerase DNA SEQ ID NO:521.
XX
KW   Eucalyptus grandis; pinus radiata; Monterey pine; modification;
KW   plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
KW   transgenic plant; ds.
XX
OS   Pinus radiata.
XX
PN   WO200022092-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999; 99WO-N200169.
XX
PR   13-OCT-1998; 98US-0170862.
XX
PR   11-AUG-1999; 99US-0148426.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
PA   (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI   Bloksberg LN;
XX
DR   MPI; 2000-339328/29.
XX
PT   New genes encoding proteins involved in a plant polysaccharide
PT   biosynthetic pathway, useful for modulating or altering the
PT   polysaccharide content, composition or structure of the plant
XX
PS   Claim 1; Page 206; 301pp; English.
XX
CC   The present invention describes isolated polynucleotides (PN) comprising
CC   a sequence selected from one of 835 nucleotide sequences given in
CC   AAA67073 to AAA67907, their (reverse) complements, sequences producing
CC   an expectation (E) value of 0.01 or less compared to the 835 sequences,
CC   sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
CC   835 sequences or sequences that are degenerately equivalent or allelic
CC   to the 835 sequences. The polynucleotides are used to modify the
CC   activity of a polypeptide involved in a polysaccharide biosynthetic
CC   pathway in the plant. They are especially used to modulate or alter the
CC   polysaccharide content, composition or structure of the plant. AAB16268
CC   to AAB16340 are proteins encoded by some of the polynucleotide sequence
CC   given in the present invention.
XX
SQ   Sequence 410 BP; 116 A; 80 C; 110 G; 104 T; 0 other;
Query Match      54.5%; Score 211.6; DB 21; Length 410;
Best Local Similarity 73.2%; Pred. No. 2.1e-57;
Matches 271; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY   9 aaagaaagatgtggaatgacacgcgaaataagctctcgtatgctctctccgactgcgc 68
DB   41 aacacagagaaggggtatcccttaaatgcccgcgcacatgtgtcatcagaacttggc 100
QY   69 aattgcttcgaggtgcgtgcgcatgtccacttcgagcgacgattgtgtccacatggac 128
DB   101 aatctgcttcagagcgcaaatatataatgacgaaatgtgtcagattgtgtcattgac 160
QY   129 atcatgattggcatttgcctcccaattacattatgctcagctcagtttgaagtgtg 188
DB   161 atcatgattggcatttgcctcccaattacattatgctcagctcagtttgaagtgtg 220
QY   189 agaaagacacacagagcatattgtgattgtcacattatggttacaatcctctgttat 248
      ||||||| || ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
```

```
DB   221 aggaagcataccagcattcttgatgtcatcttatgtacacaaacctcttgattat 280
QY   249 gttagaccttggcaaaagctgtctcctgtttatattcactcagtagacatcaaaa 308
      || || || || || || || || || || || || || || || || || || || || ||
DB   281 gtggaacatttgcgaagacttgagcttcagcttccacttctcattgtgagctgcacaa 340
QY   309 gataactggaagaacttaccgaagaatcagacatgcatgatctcctggtgtgca 368
      || || || || || || || || || || || || || || || || || || || || ||
DB   341 gacatttggcaagatctcatcacaagaatcagaatctgcgactgcgcttgagtgca 400
QY   369 ttaagcctg 378
      || || || || || || || || || || || || || || || || || || || || ||
DB   401 gtgaaacctg 410
RESULT 9
AAA67513
ID   AAA67513 standard; DNA; 502 BP.
XX
AC   AAA67513;
XX
DT   31-OCT-2000 (first entry)
XX
DE   Pinus radiata ribulose-phosphate-3-epimerase DNA SEQ ID NO:514.
XX
KW   Eucalyptus grandis; pinus radiata; Monterey pine; modification;
KW   plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
KW   transgenic plant; ds.
XX
OS   Pinus radiata.
XX
PN   WO200022092-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999; 99WO-N200169.
XX
PR   13-OCT-1998; 98US-0170862.
XX
PR   11-AUG-1999; 99US-0148426.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
PA   (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI   Bloksberg LN;
XX
DR   MPI; 2000-339328/29.
XX
PT   New genes encoding proteins involved in a plant polysaccharide
PT   biosynthetic pathway, useful for modulating or altering the
PT   polysaccharide content, composition or structure of the plant
XX
PS   Claim 1; Page 204; 301pp; English.
XX
CC   The present invention describes isolated polynucleotides (PN) comprising
CC   a sequence selected from one of 835 nucleotide sequences given in
CC   AAA67073 to AAA67907, their (reverse) complements, sequences producing
CC   an expectation (E) value of 0.01 or less compared to the 835 sequences,
CC   sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
CC   835 sequences or sequences that are degenerately equivalent or allelic
CC   to the 835 sequences. The polynucleotides are used to modify the
CC   activity of a polypeptide involved in a polysaccharide biosynthetic
CC   pathway in the plant. They are especially used to modulate or alter the
CC   polysaccharide content, composition or structure of the plant. AAB16268
CC   to AAB16340 are proteins encoded by some of the polynucleotide sequence
CC   given in the present invention.
XX
SQ   Sequence 502 BP; 138 A; 99 C; 138 G; 127 T; 0 other;
Query Match      53.6%; Score 208; DB 21; Length 502;
Best Local Similarity 72.8%; Pred. No. 3.3e-56;
Matches 268; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
```


PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match	52.6%	Score 204.2;	DB 21;	Length 620;
Best Local Similarity	74.0%	Pred. No. 6e-55;		
Matches 276; Conservative	0;	Mismatches 88;	Indels 9;	Cross 1

QY	25	tgacacccgaataatagctcccttcgattctctctccgacctcgccaattgcttcgaagg	84
Db	5	tgctgcgcggaanaatagacacgctcgatgctgtctcatcgacttcgcaatccaagcggcggaagg	64
QY	85	ctcaaggcatgctcccaactcgcgcgcgcgacttgctccacatggacatcattgattg-----	136
Db	65	ccaactcggatgacgcgatttgggcgcgaactcgtctccacatgatataatagtacgcgtcccc	120
QY	139	---ggcattttgcccccaattcaactattggcgcgcccgatattgaaagcttggaaagc	190
Db	125	gcagcgattttgtccctcaatccttaacgatttggctccctgctacatcgaggatttggaaagc	180
QY	196	acacaaagacatatttgatctgacactatggtatgtataaccccttgattatgttgtaac	250
Db	185	acactaatgcatactctgatttgccacttatggtagaaccaccatgattatgcgtgcac	240
QY	256	cccttggcacaagctgtgctcttcgtcttgaatttcaattcgaatgtagacataccaagaataact	315
Db	245	agaatggtctaaagcttgggctctctgcttcacatccacgctgaggtggcccaagataatt	304
QY	316	ggaagaagaaatttcacaagaattccaagctcaactgtgcatagtatccctgtgtgatccataaagc	375
Db	305	ggcaacaactctgtcgaagaagatcaagctcactctgtagtagaacccagcgcgtgctcttaagc	364
QY	376	ctggagaccccgct 388	
Db	365	ctggaaacacccct 377	

RESULT	11
AAC38028	
ID	AAC38028 standard; DNA; 1116 BP

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 19528
XY

Hybridisation assay; genetic mapping; gene expression control; protein identification; protein expression

metabolic pathway; promoter; termination sequence; ss.

OS *Arabidopsis thaliana*

PN EP1.033405-A2

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 09-MAR-1999; 99US-0123548

PR	25-MAR-1999:	99US-012626A
LA	23-MAR-1999:	99US-0125788.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135529.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154479.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

Sequence similarity	76.0%;	Pred No. 8.4e-39;				
Matches	190;	Conservative	0; Mismatches	60; Indels	0; Gaps	0;

RESULT 12
AAC38251 ID AAC38251 standard; DNA; 487 BP.
XX AC AAC38251;
XX DT 17-OCT-2000 (first entry)
XX DE Zea mays DNA fragment SEQ ID NO: 20310.
XX DE Hybridisation assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway; metabolic;
KM pathway; promoter; termination sequence; corn; ss.
XX XX
OS Zea mays subsp. mays.
XX
XX EPI033405-A2.
FN
EX

PD	06-SEP-2000.	
PE		
FX	25-FEB-2000;	2000EP-0301439
XX		
PR	25-FEB-1999;	990US-0121925
PR	05-MAR-1999;	990US-0123180
PR	09-MAR-1999;	990US-0123548
PR	23-MAR-1999;	990US-0125788
PR	25-MAR-1999;	990US-0126264
PR	25-MAR-1999;	990US-0126785
PR	01-APR-1999;	990US-0127462
PR	06-APR-1999;	990US-0128234
PR	08-APR-1999;	990US-0128714
PR	11-APR-1999;	990US-0129845
PR	21-APR-1999;	990US-0130077
PR	23-APR-1999;	990US-0130449
PR	28-APR-1999;	990US-0130510
PR	30-APR-1999;	990US-0130891
PR	03-MAY-1999;	990US-0131449
PR	04-MAY-1999;	990US-0132048
PR	05-MAY-1999;	990US-01332407
PR	06-MAY-1999;	990US-0133488
PR	06-MAY-1999;	990US-0132466
PR	07-MAY-1999;	990US-0132487
PR	11-MAY-1999;	990US-0132883
PR	14-MAY-1999;	990US-0134256
PR	14-MAY-1999;	990US-0134218
PR	14-MAY-1999;	990US-0134219
PR	14-MAY-1999;	990US-0134221
PR	14-MAY-1999;	990US-0134221
PR	18-MAY-1999;	990US-0134370
PR	19-MAY-1999;	990US-0134768
PR	20-MAY-1999;	990US-0134941
PR	21-MAY-1999;	990US-0135124
PR	24-MAY-1999;	990US-0135353
PR	25-MAY-1999;	990US-0135529
PR	27-MAY-1999;	990US-0136021
PR	28-MAY-1999;	990US-0136782
PR	01-JUN-1999;	990US-0136788
PR	03-JUN-1999;	990US-0137222
PR	04-JUN-1999;	990US-0137328
PR	07-JUN-1999;	990US-0137724
PR	08-JUN-1999;	990US-0137924
PR	10-JUN-1999;	990US-0138540
PR	14-JUN-1999;	990US-0138847
PR	16-JUN-1999;	990US-0139119
PR	16-JUN-1999;	990US-0139455
PR	17-JUN-1999;	990US-0138455
PR	18-JUN-1999;	990US-0138459
PR	18-JUN-1999;	990US-0139455
PR	18-JUN-1999;	990US-0139456
PR	18-JUN-1999;	990US-0139461
PR	18-JUN-1999;	990US-0139462
PR	18-JUN-1999;	990US-0139463
PR	18-JUN-1999;	990US-0139750
PR	18-JUN-1999;	990US-0139763
PR	21-JUN-1999;	990US-0139817
PR	22-JUN-1999;	990US-0139899
PR	23-JUN-1999;	990US-0140353
PR	24-JUN-1999;	990US-0140354
PR	28-JUN-1999;	990US-0140695
PR	29-JUN-1999;	990US-0140923
PR	30-JUN-1999;	990US-0140981
PR	01-JUL-1999;	990US-0141287
PR	01-JUL-1999;	990US-0141842
PR	02-JUL-1999;	990US-0142054
PR	02-JUL-1999;	990US-0142055

Db 478 caggtactcc 487

RESULT 13

AA158727 standard; cDNA; 907 bp.
XX

AC AA158727;

DT	22-OCT-2001	(first entry)
yy		

Human polynucleotide SEQ ID NO 930.

KW immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukemia; ss.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0750000.

PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0598043.

U3-AUG-2000; 2000US-0653450.
14-SEP-2000: 2000US-0663161

PR 19-NOV-2000: 2000TS-0737244

(HYSE-) HYSEO INC.

Tang YT, Llu C, Asundi V.

Zhao QA, Zhou P, Goodrich

DOI: 10.1002/wpi.2001-442253/47.

XX

such as central nervous system

Claim 1; SEQ ID NO 930; 1007

The invention relates to human

Immunosuppressant and cytostatic

of the invention may be used

localised neuropathies and ce

lateral sclerosis, and Shy-Dr

Activin/inhibin activity, cholesterol, and thrombolytic activity

assays for receptor activity, CNS disorders

CC polypeptide which is useful for treating a disease or condition
 CC associated with decreased or increased expression of functional human
 CC isomerase. ISOM polypeptides are useful for treating immune disorders
 CC such as inflammation, acquired immunodeficiency syndrome (AIDS),
 CC Addison's disease, adult respiratory distress syndrome (ARDS),
 CC allergies, anemia, asthma, atherosclerosis, Crohn's disease, atopic
 CC dermatitis, glomerulonephritis, Grave's disease, Hashimoto's
 CC thyroiditis, multiple sclerosis, osteoporosis, rheumatoid arthritis,
 CC psoriasis, myocardial or pericardial inflammation, and autoimmune
 CC diseases, diabetes mellitus, cancers and cell proliferative
 CC disorders such as actinic keratosis, prostatitis, cirrhosis, and
 CC myelofibrosis.

CC Sequence 1205 BP; 331 A; 224 C; 291 G; 359 T; 0 other;

Query Match 22.9%; Score 88.8; DB 22; Length 1205;
 Best Local Similarity 56.9%; Pred. No. 5.6e-18;
 Matches 206; Conservative 0; Mismatches 147; Indels 9; Gaps 2;

QY 33 aaaaatagctccttcgagctcctccttcgagctcgaattgtgctcgaagctcagcgc 92
 DB 21 aagattggccgcgtccatcctcaacagcagcctggccaatttaaggcgcgagtgctccgcg 80
 QY 93 atgctccacttggcgccgagtgctccacatgagacatgagatggcatttgcctcc 152
 DB 81 atgctagactcggcgccgagtgatctacacctggacgtatgagcggcatttgcctcc 140
 QY 153 aatttaactatcggcgctcagttatgaaagttgaaagca-----cacaagagca 206
 DB 141 aacatcacttctgtacacctgtgtgtagaagaagccttcgaagaagcagcttaagccagacct 200
 QY 207 tatttgagttgcacctatggttacaacatcctctgattatgttgaaccttggcaaaa 266
 DB 201 tctcttgacatcgacatgtagtgggtccaaagcagaacagtggtgtaagccaatgctgtga 260
 QY 267 gctgtggtcctcggtttacatttcacgtagagacatcaaaagatactcggaaagacct 326
 DB 261 gcaaggagcaatcagacacattcctcgcag--gctactgagaaccacagggccttg 317
 QY 327 atccaaagatcaagtcacatgagcatgctcctggtgtagtaagcctggagaccgcc 386
 DB 318 attaagacatcgcggagagatggatgaaggttgccttcgcatcaaacaggaaacctca 377

QY 387 gt 388
 DB 378 gt 379

RESULT 15

AA160513 standard; CDNA; 2831 BP.

AA160513;

22-OCT-2001 (first entry)

Human polynucleotide SEQ ID NO 4502.

Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
 peripheral nervous system; neuropathy; central nervous system; CNS;
 Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 chemokine; thrombolytic; drug screening; arthritis; inflammation;
 leukaemia; ss.

OS Homo sapiens.
 XX WO200153312-A1.
 XX 26-JUL-2001.
 PD 26-DEC-2000; 2000WO-US34263.
 XX PF

XX 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.

XX (HSE-) HSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

DR WPI: 2001-442253/47.
 DR P-PSDB; AAA41357.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -

XX Claim 1; SEQ ID NO 4502; 10078bp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA158642-AA162213) with noctropic,
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 2831 BP; 873 A; 476 C; 583 G; 897 T; 2 other;

Query Match 22.9%; Score 88.8; DB 22; Length 2831;
 Best Local Similarity 56.9%; Pred. No. 8.5e-18;
 Matches 206; Conservative 0; Mismatches 147; Indels 9; Gaps 2;

QY 33 aaaaatagctccttcgagctcctccttcgagctcgaattgtgctcgaagctcagcgc 92
 DB 73 aagattggccgcgtccatcctcaacagcagcctggccaatttaaggcgcgagtgctccgcg 132
 QY 93 atgctccacttggcgccgagtgctccacatgagacatgagatggcatttgcctcc 152
 DB 133 atgctagactcggcgccgagtgatctacacctggacgtatgagcggcatttgcctcc 192
 QY 153 aatttaactatcggcgctcagttatgaaagttgaaagca-----cacaagagca 206
 DB 193 aacatcacttggtagacctgtgtagaagaagccttcgaagaagcagtagcagagacct 252
 QY 207 tatttgagttgcacctatggttacaacatcctctgattatgttgaaccttggcaaaa 266
 DB 253 tctcttgacatcgacatgtagtgggtccaaagcagaacagtggtgtaagccaatgctgtga 312
 QY 267 gctgtggtcctcggtttacatttcacgtagagacatcaaaagatactcggaaagacct 326
 DB 313 gcaaggagcaatcagacacattcctcgcag--gctactgagaaccacagggccttg 369
 QY 327 atccaaagatcaagtcacatgagcatgctcctggtgtagtaagcctggagaccgcc 386
 DB 370 attaagacatcgcggagagatggatgaaggttgccttcgcatcaaacaggaaacctca 429
 QY 387 gt 388

Wed Jul 3 12:01:05 2002

us-09-300-482-298.rng

Page 18

Db 430 gt 431

Search completed: July 3, 2002, 07:57:43
Job time: 24418 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 07:56:01 ; Search time 185.98 Seconds
(without alignments)
512.452 Million cell updates/sec

Title: US-09-300-482-298

Perfect score: 388

Sequence: 1 ggagaaagaagaagaatg.....ttaagcctggaccacctt 388

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCRTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88.8	22.9	3556	4	US-09-276-531-8
2	65.4	16.9	1234	4	US-08-858-207A-110
3	58	14.9	1019	3	US-08-714-918-79
4	58	14.9	1019	4	US-09-265-315-79
5	58	14.9	1019	4	US-09-265-315-79
6	58	14.9	1019	4	US-09-265-315-79
7	34	8.8	3781	3	US-08-688-988-5
8	31.6	8.1	40352	3	US-08-846-111D-15
9	31	8.0	1413	4	US-08-900-117A-3
10	31	8.0	1430	2	US-08-276-452A-25
11	31	8.0	1430	2	US-08-798-744-25
12	30.8	7.9	908	4	US-09-457-046B-17
13	30.6	7.9	3150	3	US-08-673-814-5
14	30.6	7.9	3150	4	US-09-115-824-5
15	30.2	7.8	4411529	4	US-09-103-840A-1
16	29.8	7.7	2168	3	US-08-749-522-6
17	29.8	7.7	5300	1	US-08-766-014-23
18	29.8	7.7	5300	1	US-08-766-014-23
19	29.8	7.7	5300	1	US-08-766-014-23
20	29.6	7.6	911	4	US-09-813-817-3
21	29.6	7.6	911	4	US-09-457-046B-19
22	29.6	7.6	1326	4	US-09-457-046B-21
23	29.4	7.6	1474	4	US-08-821-994-64
24	29.4	7.6	6960	2	US-08-841-349-3
25	29.4	7.6	8176	2	US-08-841-349-3
26	29	7.5	4182	2	US-07-973-257-1
27	29	7.5	5181	1	US-08-257-073-10

C 28	29	7.5	8700	2	US-08-392-625-16	Sequence 16, Appl
C 29	29	7.5	8700	2	US-08-466-961A-16	Sequence 16, Appl
C 30	28.6	7.4	15872	4	US-09-105-537-1	Sequence 1, Appl
C 31	28.4	7.3	1347	3	US-08-663-713A-1	Sequence 1, Appl
C 32	28.4	7.3	1347	3	US-09-014-888-1	Sequence 2, Appl
C 33	28.4	7.3	2835	1	US-08-750-533-2	Sequence 2, Appl
C 34	28.4	7.3	4765	1	US-08-750-533-8	Sequence 8, Appl
C 35	28.4	7.3	4765	4	US-08-894-818B-7	Sequence 7, Appl
C 36	28.4	7.3	4765	4	US-09-445-472-5	Sequence 5, Appl
C 37	28.4	7.3	10207	1	US-08-920-812-2	Sequence 2, Appl
C 38	28.4	7.3	10207	1	US-08-920-812-2	Sequence 2, Appl
C 39	28.4	7.3	10207	1	US-08-921-177-2	Sequence 2, Appl
C 40	28.4	7.3	10207	1	US-08-362-577C-2	Sequence 2, Appl
C 41	28.4	7.3	10207	2	US-08-920-828-2	Sequence 2, Appl
C 42	28.2	7.3	305	4	US-08-821-994-81	Sequence 81, Appl
C 43	28.2	7.3	753	4	US-08-861-774E-71	Sequence 71, Appl
C 44	28.2	7.3	924	4	US-09-105-390-33	Sequence 33, Appl
C 45	28.2	7.3	1035	4	US-09-105-390-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-09-276-531-8
Sequence 8, Application US/09276531
Patent No. 6183968
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hallman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Reddy, Roopa
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
TITLE OF INVENTION: RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276,531
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/079,677
FILING DATE: March 27, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Lynn E. Murty, Ph.D.
REGISTRATION NUMBER: 42,918
REFERENCE/DOCKET NUMBER: PA-0008 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3556 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LIVERUT01

CLONE: 1753826
US-09-276-531-8

Query Match
Best Local Similarity 22.9%; Score 88.8; DB 4; Length 3556;
Matches 206; Conservative 0; Mismatches 147; Indels 9; Gaps 2;

QY 33 aaatagctccttcgaagctcttcgcagcttcgcaatttgctcgaagctcagcgc 92
DB 72 AAGATTGGCCCGTCATCCCTCAACAGGACCTGGCCCAATTAGGGGCGAGTGCCTCCGG 131
QY 93 atgtccacttcgagcgagcttcgctccacagagacatcagatgagggcatttgcctcc 152
DB 132 ATGCTTAACCTCTGGGCGCATTTATCTGCACCTGGAGTAATGACGGCATTGTTGCC 191
QY 153 aatttaactatctggcgctcagcttattgaaagttgaaagaaagca-----cacaagagca 206
DB 192 AACATCACCCTTGGTCACCTGTGGTAAGAACCTTGCAGAAAGCAGCTAGGCCAGACCTT 251
QY 207 tatttgacttcacactatggtttcaaaatccctctgtatattgtgaaccttgagcaaa 266
DB 252 TCTTTGACATGCACATGATGTGTGCAAGCCAGACAGTGGTAAGCCATGGCTGTA 311
QY 267 gctgtgctctcgtgtttacattcaacgtagagacatcaaaagtaacttgaaagact 326
DB 312 GAGAGAGCCAAATCACTACACCTTTCATCTCGAG---GCTACTGAGAACCCAGGGCTTGG 368
QY 327 atccaaagaatcaagtaacatgcatgatactcctgtgtgtagcatlaagcctggagacccc 386
DB 369 ATTAAAGACATTCGGAGAAATGAGATGAAGTTGGCCTTCATCAAAACGAGAACTCTCA 428
QY 387 GT 388
DB 429 GT 430

RESULT 2

US-08-858-207A-110
Sequence 110, Application US/08858207A
Patent No. 6348328

GENERAL INFORMATION:

APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Nicholas, Richard
APPLICANT: Stodola, Robert
TITLE OF INVENTION: No. 6348328el Compounds
NUMBER OF SEQUENCES: 552
CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA

COUNTRY: USA
ZIP: 19406-0939

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/858,207A

FILING DATE: 09-MAY-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/017670

FILING DATE: 14-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Gimm, Edward R

REGISTRATION NUMBER: 38,891

REFERENCE/DOCKET NUMBER: P50475

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090

INFORMATION FOR SEQ ID NO: 110:

SEQUENCE CHARACTERISTICS:

LENGTH: 1234 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-858-207A-110

Query Match
Best Local Similarity 16.9%; Score 65.4; DB 4; Length 1234;
Matches 126; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 33 aaatagctccttcgaagctcttcgcagcttcgcaatttgctcgaagctcagcgc 92
DB 956 AAGATTGGCCCGTCATCCCTCAACAGGACCTGGCCCAATTAGGGGCGAGTGCCTCCGG 131
QY 93 atgtccacttcgagcgagcttcgctccacagagacatcagatgagggcatttgcctcc 152
DB 1016 CTAGAACCACTGGGCGCAATATGCGCAGATATGCCCATATGATATGATGACAGTCAATTTGTACCG 1075
QY 153 aatttaactatctggcgctcagcttattgaaagttgaaagaaagcaacaaagacatttg 212
DB 1076 CAATCAGTTTGTGTGTCAGTGTGTGTCAGACCTTCTCTCATAGTAAGATGATGTTTC 1135
QY 213 gatgtgacttcgtttcaaaatccctctgtatattgtgaacctt 259
DB 1136 GATTGCACCTTGTGTGTGTCAGAACCTTGCATCATCTGGAATAATT 1182

RESULT 3

US-08-714-918-79/c
Sequence 79, Application US/08714918
Patent No. 6037123

GENERAL INFORMATION:

APPLICANT: Benton, Bret
APPLICANT: Lee, Ving
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/714,918

FILING DATE: September 13, 1996

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,102

FILING DATE: December 22, 1995

APPLICATION NUMBER: 60/003,798

FILING DATE: September 15, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Walburg, Richard J.

REGISTRATION NUMBER: 32,327


```

: REFERENCE/DOCKET NUMBER: 222/005
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 79:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1019 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-714-918-79

Query Match
Best Local Similarity 56.9%; Score 58; DB 3; Length 1019;
Matches 161; Conservative 0; Mismatches 118; Indels 4; Gaps 3;

QY 105 ggcgcgattggctccacatgacatcatgagatggcaattgtcccaattactatt 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1001 GGNITCGACGAGTTCATTTTGATGATGATGATGATGATGATGATGATGATGAT 942

QY 165 ggcgcctccagttatgaaattgaaagacacacaaagcatatttgatgtccactt 224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 941 GGTTCACCAATATTAGATGCGATGAAGCAACACATTACTATATAGACGTACATTTG 882

QY 225 atggtt-acaatcctcttgatattgtgaaccttgcaaaagctgtgctct-9gtt 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 881 ATGATTGAATAATTCAGAAAGTATATTGCATCATTTGCACACATGTCGCGATATGAT 822

QY 283 ttacattcagtagagacatcaaaagtaactggaagaactatccaaagaactcaagt 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 821 TTNCAATTCAATGTCGAATCAACGCCCTCATTA--TTTCATCGTGTCTATTCAAATGATTTAAC 764

QY 343 cacatggcatgatcctcggtgtagcatlaaagcctgggacccc 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 763 ATTAGATTAATAAAGCTGCTGTAGTAAATTAATCTGTGTACACC 721

RESULT 4
US-09-265-315-79/c
: Sequence 79, Application US/09265315
: Patent No. 6187541
: GENERAL INFORMATION:
: APPLICANT: Benton, Bret
: APPLICANT: Lee, Ying J.
: APPLICANT: Malouin, Francois
: APPLICANT: Martin, Patrick K.
: APPLICANT: Schmid, Molly B.
: APPLICANT: Sun, Dongxu
: TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
: TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
: TITLE OF INVENTION: TARGET GENES
: NUMBER OF SEQUENCES: 111
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071-2066
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 MB
: MEDIUM TYPE: storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: Word Perfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/265,315
: FILING DATE: March 9, 1999
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
```

```

: APPLICATION NUMBER: 08/714,918
: FILING DATE: September 13, 1996
: APPLICATION NUMBER: 60/009,102
: FILING DATE: December 22, 1995
: APPLICATION NUMBER: 60/003,798
: FILING DATE: September 15, 1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Warburg, Richard J.
: REGISTRATION NUMBER: 32,327
: REFERENCE/DOCKET NUMBER: 240/247
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 79:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1019 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-09-265-315-79

Query Match
Best Local Similarity 56.9%; Score 58; DB 4; Length 1019;
Matches 161; Conservative 0; Mismatches 118; Indels 4; Gaps 3;

QY 105 ggcgcgattggctccacatgacatcatgagatggcaattgtcccaattactatt 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1001 GGNITCGACGAGTTCATTTTGATGATGATGATGATGATGATGATGATGATGAT 942

QY 165 ggcgcctccagttatgaaattgaaagacacacaaagcatatttgatgtccactt 224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 941 GGTTCACCAATATTAGATGCGATGAAGCAACACATTACTATATAGACGTACATTTG 882

QY 225 atggtt-acaatcctcttgatattgtgaaccttgcaaaagctgtgctct-9gtt 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 821 TTNCAATTCAATGTCGAATCAACGCCCTCATTA--TTTCATCGTGTCTATTCAAATGATTTAAC 764

QY 283 ttacattcagtagagacatcaaaagtaactggaagaactatccaaagaactcaagt 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 821 TTNCAATTCAATGTCGAATCAACGCCCTCATTA--TTTCATCGTGTCTATTCAAATGATTTAAC 764

QY 343 cacatggcatgatcctcggtgtagcatlaaagcctgggacccc 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 763 ATTAGATTAATAAAGCTGCTGTAGTAAATTAATCTGTGTACACC 721

RESULT 5
US-09-265-315-79/c
: Sequence 79, Application US/09265315
: Patent No. 6187541
: GENERAL INFORMATION:
: APPLICANT: Benton, Bret
: APPLICANT: Lee, Ying J.
: APPLICANT: Malouin, Francois
: APPLICANT: Martin, Patrick K.
: APPLICANT: Schmid, Molly B.
: APPLICANT: Sun, Dongxu
: TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
: TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
: TITLE OF INVENTION: TARGET GENES
: NUMBER OF SEQUENCES: 111
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071-2066
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 MB
```



```

MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/265,315
FILING DATE: March 9, 1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/714,918
FILING DATE: September 13, 1996
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Marburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 240/247
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 1019 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-265-315-79

```

```

Query Match      14.9%  Score 58; DB 4; Length 1019;
Best Local Similarity 56.9%; Pred. No. 5.8e-10;
Matches 161; Conservative 0; Mismatches 118; Indels 4; Gaps 3;

QY 105 ggcgcgaattggcccaatgacatcatgatggagcatttgcctcccaatttaactt 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1001 GGNTCGACGAGAGTTTGAATTTGATGATGATGATGATGATGATGATGATGAT 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 165 ggcgcgaatttgaagtltgagaagcacacaaggcatatttgatgtcaactt 224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 941 GGTTCACCAATATTGATGACAGTAAGAAAGCACAACATTACNTATGACGTACATTTG 882
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 225 atggct-acaacaccttgatattgtaaaccttggcaaaagctggtctct-99tt 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 881 ATGATTGAAAATTCAGAAAAGTATATTCATCATTTGCAGAACATGTCGCGATATGAT 822
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 283 ttacattcaagtagagacatcaaaagataactggaagaacttatccaaagaatcaagt 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 821 TTNCATTTGATGTCGAATCAACGCCCTCATTA--TTTCATCGTGTCTATTCAATGATTAAAC 764
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 343 cacatgacatgattcctgtgtgtagatctaaagcctggagccccc 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 763 ATTTAGATTAATAAAGCTGTGTGATTAATTAATTCCTGTACACC 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 6
US-09-266-417-79/c
Sequence 79, Application US/09266417
Patent No. 6228588
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving J.
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:

```

```

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/266,417
FILING DATE: March 9, 1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/714,918
FILING DATE: September 13, 1996
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Marburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 240/248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 1019 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-266-417-79

```

```

Query Match      14.9%  Score 58; DB 4; Length 1019;
Best Local Similarity 56.9%; Pred. No. 5.8e-10;
Matches 161; Conservative 0; Mismatches 118; Indels 4; Gaps 3;

QY 105 ggcgcgaattggcccaatgacatcatgatggagcatttgcctcccaatttaactt 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1001 GGNTCGACGAGAGTTTGAATTTGATGATGATGATGATGATGATGATGATGAT 164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 165 ggcgcgaatttgaagtltgagaagcacacaaggcatatttgatgtcaactt 224
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 941 GGTTCACCAATATTGATGACAGTAAGAAAGCACAACATTACNTATGACGTACATTTG 882
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 225 atggct-acaacaccttgatattgtaaaccttggcaaaagctggtctct-99tt 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 881 ATGATTGAAAATTCAGAAAAGTATATTCATCATTTGCAGAAACATGTCGCGATATGAT 822
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 283 ttacattcaagtagagacatcaaaagataactggaagaacttatccaaagaatcaagt 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 821 TTNCATTTGATGTCGAATCAACGCCCTCATTA--TTTCATCGTGTCTATTCAATGATTAAAC 764
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 343 cacatgacatgattcctgtgtgtagatctaaagcctggagccccc 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 763 ATTTAGATTAATAAAGCTGTGTGATTAATTAATTCCTGTACACC 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 7
US-08-688-988-5/c
Sequence 5, Application US/08688988B
Patent No. 6096545
GENERAL INFORMATION:
APPLICANT: Lefebvre, Daniel D.
APPLICANT: Malboobi, Mohammad A.

```


QY	127	acacatgagatggagcatttggccccaatttaaccttgagcgctccagttatgaaagt	186
Db	1678	AAATAAACGGATTAAAGCTTTTAAACCAACTTTTCTGTAAACCGGTTTTGTAGTTT	1611
QY	187	tggaaagcacacaagaagcaatttgatgtgcacttggttacaactcctctgatt	246
Db	1518	TTCTGTCGGACTGGCCCTAATAAGTTGTTCCATATTTATCCCTTAATAATTCATATATT	1555
QY	247	atgctgaaccccttgcanaag	267
Db	1558	AATAGGCACCAATAGCATTTAG	1538

```

RESULT 15
: Sequence 1, Application US/09103840A
: Patent No. 6294328
: GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: FILE REFERENCE: 24366-20007.00
: CURRENT APPLICATION NUMBER: US/09/103,840A
: NUMBER OF SEQ. ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.1
: SEQ. ID NO 1
: LENGTH: 4411529
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: OTHER INFORMATION: H37Rv
: US-09-103-840A-1

```

[illegible]

Search completed: July 3, 2002, 08:04:53
Job time: 25458 sec

Query Match	7.9%	Score 30.6;	DB 4;	Length 3150;
Best Local Similarity	51.1%;	Pred. No. 2.2;		
Matches	72;	Conservative	0;	Mismatches 69;
			Indels	0;
			Gaps	0

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 05:12:38 ; Search time 5254.53 Seconds

(without alignments)
1091.224 Million cell updates/sec

Title: US-09-300-482-311

Perfect score: 274
Sequence: 1 ctcatcctcttctccacc.....gccaaagatgcgcgtgcacg 274

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GeneBld:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_ph:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description

1	81.4	29.7	798	8	AY066038	AY066038 Arabidops
2	81.4	29.7	959	8	AF360330	AF360330 Arabidops
3	81.4	29.7	1464	8	AY054172	AY054172 Arabidops
4	81.4	29.7	66653	8	AC016972	AC016972 Arabidops
5	81.4	29.7	84957	8	AC006200	AC006200 Arabidops
6	68	24.8	1118	6	A93589	A93589 Sequence 1
7	68	24.8	1118	6	SPR5PR	SPR5PR
8	62	22.6	1084	8	AF154658	AF154658 Nicotiana
9	54.6	19.9	135296	2	AP003944	AP003944 Oryza sat
10	54.6	19.9	136775	2	AP004279	AP004279 Oryza sat
11	54.6	19.9	150638	2	AP003616	AP003616 Oryza sat
12	54	19.7	998	8	AY045785	AY045785 Arabidops
13	54	19.7	93234	8	AC009465	AC009465 Arabidops
14	54	19.7	95310	8	ATAC011437	ATAC011437 Arabidops
15	49.2	18.0	454	6	AX172623	AX172623 Sequence
16	48.4	17.7	34158	1	SCR13	SCR13
17	47.6	17.4	152396	8	AP002971	AP002971 Oryza sat
18	47.6	17.4	163089	2	AP004358	AP004358 Oryza sat
19	47.4	17.3	37898	1	SCE34	SCE34
20	46.6	17.0	11727	1	AE008449	AE008449 Streptoco
21	46.6	17.0	36224	1	SCD78	SCD78
22	46.6	17.0	144561	2	AP004307	AP004307 Oryza sat
23	46.6	17.0	297172	2	SPNEU1905	SPNEU1905
24	46.4	16.9	1288	6	AR007563	AR007563 Sequence
25	46.4	16.9	161266	8	AR001278	AR001278 Oryza sat
26	45.8	16.7	3800	1	SAR416377	AJ416377 Streptomy
27	45.8	16.7	14916	1	AE005900	AE005900 Caulobact
28	45	16.4	10834	1	AE004020	AE004020 Xylella f
29	44.8	16.4	3875	1	RRPHAS	X66407 R. ruber ORF
30	44.6	16.3	15424	1	AE007019	AE007019 Mycobacte
31	44.6	16.3	40790	1	MTC1493	295844 Mycobacteri
32	44.6	16.3	77534	1	AF235504	AF235504 Streptomy
33	44.6	16.3	139399	8	AP002865	AP002865 Oryza sat
34	44.6	16.3	182756	8	AC007789	AC007789 Oryza sat
35	44.4	16.2	109990	1	AE005969	AE005969 Caulobact
36	44.4	16.2	43632	1	SC6A5	AL049485 Streptomy
37	44.4	16.2	190097	2	AC094543	AC094543 Rattus no
38	44.2	16.1	1593	1	AF113605	AF113605 Streptomy
39	44	16.1	2647	8	OSHPRG	X61280 O. sativa ge
40	44	16.1	28732	1	AF082100	AF082100 Streptomy
41	43.8	16.0	273	1	MTU13938	U13938 Mycobacteri
42	43.8	16.0	194415	2	AP004365	AP004365 Oryza sat
43	43.6	15.9	42273	8	NCB12N19	AL669987 Neurospor
44	43.6	15.9	69943	8	NCB11E6	AL442043 Neurospor
45	43.6	15.9	154561	8	AP002819	AP002819 Oryza sat

ALIGNMENTS

RESULT 1	AY066038	798 bp	mrna	linear	PLN 26-DEC-2001
LOCUS	Arabidopsis thaliana	At2g01290/F10A8.17	mrna, complete cds.		
DEFINITION	AY066038				
ACCESSION	AY066038.1	GI:17978866			
VERSION					
KEYWORDS	FLI CDNA.				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Kim,C.J., Chen,H., Cheuk,R., Meyers,M.C., Banh,J., Bowser,L., Carinci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayshtaki,Y., Ishida,J., Jones,T., Kamuya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.				
AUTHORS	Kim,C.J., Chen,H., Cheuk,R., Meyers,M.C., Banh,J., Bowser,L., Carinci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayshtaki,Y., Ishida,J., Jones,T., Kamuya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.				
TITLE	Arabidopsis ORF clones				
JOURNAL	Unpublished				

REFERENCE	2	(bases 1 to 998)	Kim,C.J., Chen,H., Cheuk,R., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
JOURNAL	Submitted (10-DEC-2001)	Salik Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salik Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA	
TITLE	Direct Submission		
COMMENT			RIKEN Genomic Sciences Center (GSC) members carried out the sequencing and annotation of the RAFL CDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.
FEATURES			The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL CDNAs: Kim,C.J., Chen,H., Cheuk,R., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Chang,E., Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura, Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.
source			Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.
CDS			Location/Qualifiers 1..798 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="2" /clone="013949" /note="This clone is in pUNI 51 ecotype: Columbia" 1..798 /note="putative ribose 5-phosphate isomerase" /codon_start=1 /product="At2g01290/F10A8.17" /protein_id="AA147405.1" /db_xref="GI:17978867" /translation="MALADPFLFTSDKSLAPDVAASPPQPNMLTDELKRIAYKA VEPESGAVLIGTSTAKAHAVPRIGELIRLOKLENTICIPISKTOBEALSIGPLS DUDAPVIDLSIGDAEDVPFINLYVGRGSLIREKIMIGASKRYVATVDDSKWMTI GGSKIALPVEIIPRCMKFTAKELKSLBLEGTGCEANLDEKRAVPTVDGNITVDMHV EEDMDGAGSDAILRLPGEVGEHMFDMASVIIVAGELGVKIKMKH"
BASE COUNT			193 a 183 c 227 g 195 t
ORIGIN			
Query Match			29.7%: Score 81.4; DB 8; Length 798;
Best Local Similarity			64.7%: Pred. No. 8.1e-06;
Matches 121; Conservative			0; Mismatches 66; Indels 0; Gaps 0;
QY	88	accgccatttcgcacacgcgaagaagccgcacgacgcgcctctccaccctcct 147	
DB	17	ATTCCTCTCTCATTCATCGGACAAATCTTGTCCGCTTGATGTTCCTTCACGCC 76	
QY	148	ccccctcgcgtcatctctcaaccacgaagcattggaagaaatcgcgcgtacaagccgtcg 207	
DB	77	CTCAACCCATGATTTTAACACAGACAGCTCAAAAGCTATCGCGCTTACAAAGCGGTGC 136	
QY	208	agtaagctggaagtcgcgaagtcgtcctcgacctgaagccgcgtccaccgcgaagacgtcg 267	
DB	137	AATTCGTGAGATCCGCGCATGTGTTCTCGGTCTCGAAGCCGCTCCACCGCAACACGCGC 196	
QY	268	tcgacgcg 274	

Db	197	TCGACGC 203
RESULT	2	
AF360330		
LOCUS		
DEFINITION		
AF360330		959 bp mRNA linear PLN 22-MAR-2001
Arabidopsis thaliana putative ribose 5-phosphate isomerase (F23N20.9) mRNA, complete cds.		
ACCESSION		
AF360330		
VERSION		
AF360330.1		GI:13430835
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
REFERENCE		
AUTHORS		
2 (bases 1 to 959)		
Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
Submitted (14-MAR-2001)		
Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA		
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL CDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.		
The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL CDNAs: Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Chen,H., Cheuk,R., Jones,T., Kamiya,A., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.		
Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.		
Location/Qualifiers		
1..959		
/organism="Arabidopsis thaliana"		
/db_xref="taxon:3702"		
/chromosome="1"		
/clone="RAFL09-14-p21 (R09450)"		
/note="This clone is in a modified pBluescript vector (pUC-1) as a BamHI/XhoI insert.		
ecotype: Columbia"		
1..2		
/gene="F23N20.9"		
/note="compared to genomic sequence"		
/replac="lc"		
1..26		
/gene="F23N20.9"		
1..959		
/gene="F23N20.9"		
27..830		


```
/codon_start=1
/evidence=experimental
/product="putative ribose 5-phosphate isomerase"
/protein_id="AAK26040.1"
/db_xref="GI:13430836"
/translation="MGSAPDPLVATIEDIAAVNSAPPLSNLTQELIKIAAYKAEV
ESGMVIGLGTSTAKHAAVARISLELRKGLNDIGIPSTTHEDQAVSIGPLSDIDS
HFWVLSIDGAEVDPALNLVKGSGSLREKMIKASRKFVYIDESKLVYIGSG
LAPVEVPFCCDFTRKLEELFRDSCVAKLRMKIGSGEAAVAVTDNRNVYVDLY
LERDIDGLEVAESAILRPFQVYKHGMFLGMATTLIVAGKFGTVADREC"
misc_difference
747
/gene="P23N20.9"
/note="compared to genomic sequence resulting in an amino
acid sequence difference"
/replace="g"
831..959
3'UTR
/gene="P23N20.9"
BASE COUNT 254 a 188 c 265 g 252 t
ORIGIN
```

Query Match 29.7%; Score 81.4; DB 8; Length 959;
Best Local Similarity 65.9%; Pred. No. 8e-06;
Matches 118; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 96 ttcatgcacacgagagaagccgcatgagcgcgcctcctccaccctcctcc 155
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39 TTGCATCCCTTGTCTGACGGCAGCTGAGAMCTGCGCGCTCAACTGCGCGCACCTC 98

Qy 156 gtatcctcaccacgaagcatttgaagaatactgcgcgtacacgaagcgtgagtcgtg 215
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 99 TCMAACCTCAGCGAAGAACTCAAGAAATCGCGCTTAAACCGCTGCAATTCGTC 158

Qy 216 gaatccgcatgctgctcgcgcgtacgacgcggtccacgcgcaagcattgcgtgacgcg 274
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 159 GAATCTGGAATGTTATCGCTCTCGGCACTGCTCCACCGCGAAGACCGCGTGCTCG 217

RESULT 3
AY054172 1464 bp mRNA linear PLN 30-SEP-2001
LOCUS Arabidopsis thaliana At2g01290/F10A8.17 mRNA, complete cds.
DEFINITION AY054172.1 GI:15809809
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1464)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Arabidopsis cDNA clones
TITLE
JOURNAL
REFERENCE
AUTHORS
2 (bases 1 to 1464)
Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Banh,J.,
Bower,L., Carninci,P., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y.,
Ishida,J., Jiang,P.X., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M.,
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K.,
Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C.,
Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K.,
Davis,R.W., Theologis,A. and Ecker,J.R.
Direct Submission
Submitted (22-AUG-2001) Salk Institute Genomic Analysis Laboratory
(Signal), Plant Biology Laboratory, The Salk Institute for

Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Koesema,E., Meyers,M.C., Shinn,P., Banh,J., Bower,L., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

Location/Qualifiers
1..1464
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="2"
/clone="RAFL09-23-D14(R13949)"
/note="ecotype: Columbia"
1..393
394..1191
/note="putative ribose 5-phosphate isomerase"
/codon_start=1
/product="At2g01290/F10A8.17"
/protein_id="AAU06833.1"
/db_xref="GI:15809810"
/translation="MALADPLFITSLSAFVYASPPQPMNLTOELKRIAYKA
VEFVSGMYLGLGTSTAKHAAVDRIIGELRGLKENTYIGIPSKTQEDALSGLPLS
DIDAHVVIDSLDGADEVDPFLNLVKGSGSLREKMIKASRKFVYIDDSKMKVHI
GGSKLALPVEIIPFCWKFPAEKLRSILREYGEANRLIEKAKAVTDGNVYDMHV
EDMDGDGLGAVSDAIIRLPQVEHGMFLDMSAVYIAGELGVKIKKKH"
1192..1464
3'UTR
BASE COUNT 325 a 336 c 320 g 483 t
ORIGIN

Query Match 29.7%; Score 81.4; DB 8; Length 1464;
Best Local Similarity 64.7%; Pred. No. 7.8e-06;
Matches 121; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy 88 accccaattcattcgcacacgagaagccgcgcatgagcgcgcctcctccaccctcct 147
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 ATCCCTCTTTCATTAACACGACGAGCTCAACAGTATCGCCGCTTACACCGC 469

Qy 148 ccccccgcgcattcaccacacgaagcatttgaagaatactgcgcgtacacgaagcgcgtg 207
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 470 CTAGAGCCATGAATTTAAACACAGACGAGCTCAACAGTATCGCCGCTTACAAAGCCGTG 529

Qy 208 agtacgtgagtcgcgcgcgtcgcgtacgacacgcgcgtcaccacgaagcgtgacg 267
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 530 AATTCTCGAGTCGCGCATGCTCTCGGTCGGAACCGGCTCCACCGCAACACGCGG 589

Qy 268 tcgacgcg 274
|||||
Db 590 TCGACCG 596

RESULT 4
AC016972 66653 bp DNA linear PLN 19-JAN-2001
LOCUS AC016972
DEFINITION Arabidopsis thaliana chromosome 1 BAC F23N20 genomic sequence,
complete sequence.
ACCESSION AC016972
VERSION AC016972.7 GI:12323418
KEYWORDS HTG.

Accession	Definition	LOCUS	Result
AC006200.2	Arabidopsis thaliana chromosome II section 2 of 255 of the complete sequence. Sequence from clones F23H14, F10A8.	AC006200/c	84957 bp DNA linear PLN 05-APR-2000
AC006200.1			
AC006200.2			GI:5598523

KEYWORDS	HTG.
SOURCE	thae cross.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids II: Brassicales: Brassicaceae; Arabidopsis. 1 (bases 1 to 84957)
AUTHORS	Lin X., Kaul, S., Ronnesby, S.D., Shea, T.P., Benito, M.-I., Town, C.D., Fujii, C.Y., Mason, T.M., Bowman, C.L., Barnstead, M.E., Feldblum, T.V., Buell, C.R., Ketchum, K.A., Lee, J.-J., Ronning, C.M., Koo, H., Moffatt, K.S., Cronin, L.A., Shen, M., VanAken, S.E., Unayam, L., Tallon, L.J., Gill, J.E., Adams, M.D., Carrera, A.J., Creasy, T.H., Goodman, H.M., Somerville, C.R., Copenhaver, G.P., Preuss, D., Nierman, W.C., White, O., Eisen, J.A., Salzberg, S.L., Fraser, C.M. and Venter, J.C.
TITLE	Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana
JOURNAL	Nature 402 (6763), 761-768 (1999)
MEDLINE	20083487
PUBMED	10617197
REFERENCE	2 (bases 1 to 84957)
TITLE	Lin, X.
AUTHORS	Direct Submission
COMMENT	Submitted (09-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA On Dec 17, 1999 this sequence version replaced g1:426221. The sequence and annotation of chromosome 2 were merged from those of the individual clones on this chromosome after removing overlaps. For detailed information, please see the TIGR web site (http://www.tigr.org/tdb/at/at.html).

Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (<http://artbur.epm.ornl.gov/pub/gyrail>), Genefinder (Phil Green, University of Washington), GENSCAN (Chris Burge, <http://genome.stanford.edu/GENSCAN.html>), and MetPlantgene (<http://www.cbs.dtu.dk/services/MetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curations based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by repeatmasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are numbered from the top to bottom of the chromosome.

We thank the CSHL/Masnu/ABI consortium for sequencing BAC clones F6g23, F5j6, T17A5, and T13L6, the ESSA group for sequencing clone F13D4, and Scott Jackson, Jining Jiang, Klaus Meyer, Eric Richards and Setsoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khataik, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: atetigr.org.

Location/Qualifiers

1. 84957

/organism="Arabidopsis thaliana"

/culivar="Columbia"

/db_xref="taxon:3702"

/chromosome="II"

<1. 2020

/note="Sequence from clone F23H14"


```
mRNA
complement(join(<463..924,1053..1163,1265..1417,
1513..1572,1662..1967,2287..2435,2515..2584,2718..2921,
3271..3465,3673..3945,4026..4187,4325..4455,4541..4700,
4817..5005,5145..5219))
/gene="At2g01130"
complement(<463..>5219)
/gene="At2g01130"
/note="F10A8.1"
complement(join(<463..924,1053..1163,1265..1417,1513..1572,
1662..1967,2287..2435,2515..2584,2718..2921,3271..3465,
3673..3945,4026..4187,4325..4455,4541..4700,4817..5005,
5145..5219))
/gene="At2g01130"
/codon_start=1
/product="putative Arp-dependent RNA helicase A"
/protein_id="AADI4515.2"
/db_xref="GI:6598524"
/translation="MDSRTSLPARKQSHSVLTATISQNVIVISGEGCGKTQIPQF
IIESIEANRGAFSSITCTQPRISAMSVSERVAERGEQGEISYGYVRLEGKGRD
TRLFTCTGILRLRLVDRLNGVHVIVDEIHERGMNEDLLILKDLRSRSELK
ILMSATLDAELESSYFGAGVYIIPGFYIPVRSHELEDILEMTRYRLPYNOIDXG
BETWMNKQIEPKRKSOITPVDALEADFEKPEPESRESICWYPDGIFNLIFFLIC
NICENEGGGILIFLTGMDISLSKEKQIPIIGNPDVVLVLLACHGMEFEDRLIF
EPPAGVGRVIRPGCYHLIPKCYDAFETQLPEILRTPLSLCLQIKSLNISISE
AQRGRGRRVRPGCYHLIPKCYDAFETQLPEILRTPLSLCLQIKSLNISISE
FLSRALQSPPELLAVOKATFLKIIIGALDENEDLTGLYLSKLMPEKLGKMLITGAE
LGCDPILITVAAGLSVRDPELTPQDKDLAEAKSOFSDSHDLALVRAVEGKKAE
ESAVAYDCWKNFLSIQSLRAIDSLKEFFSLIKDTGIDGNPSICEGDANLTVA
VICQMGPGICSVVHNERFSLKTMEDQVLLYNSENARETKIPYVIAFNEKIKN
SVFLADSTACSDTLILEGGSISKDGDKHLMGGLYLEPMKRDVAEITQTLKEED
ELIQKILNPVDMOAHHELISARILVSESGCGRRVFGHQLRLPILSLSKPSFL
FSRTSEGGAGNSKSOLOTLITRAGTYVPKTKIQLNKRFOTVFEFNETQIMGQPS
2021..>84957
/note="Sequence from clone F10A8"
complement(join(6441..7261,7365..7474,7570..7839,
7952..>8107))
/gene="At2g01140"
complement(6441..>8107)
/gene="At2g01140"
/note="F10A8.2"
5543..6582
/rpt_family="(CAAAA)n"
complement(join(6622..7261,7365..7474,7570..7839,
7952..8107))
/gene="At2g01140"
/codon_start=1
/product="putative fructose-bisphosphate aldolase,
plastidic form"
/protein_id="AADI4543.1"
/db_xref="GI:4262250"
/translation="MASASFVKPNLTSSPMIGQSFAMHSASSPPRYSPFARAGAY
SDLVLTAKSIASPGRIILADESNATCGKRLASITLDMEDRQAYROLTLTFELG
DIYSGLTLEETLYOSYKQSGKPFVCLDRANVYIGKIGSLPLAGSNEESWCOGLD
GLASRAEYKQCARAKRRTVYVPCGSPALVCAANGALAYAIISODNGLVPIVE
ETVADFTLMLKRRVPAVGIMFLSGGSEAEATNLNAMNOSPMPHVSYSYARAL
QNSVLRTWQKPKIEASQALKVRAKANSIAQLGKYSAGEDEDAKKMFYKQTY"
complement(9948..10066)
/rpt_family="(TAAA)n"
<12241..12943
<12241..12943
/gene="At2g01150"
<12241..12943
/gene="At2g01150"
/note="F10A8.3: identical to GP.AF078823"
12241..12762
/gene="At2g01150"
/codon_start=1
/product="RING-H2 zinc finger protein (RNA2b)"
/protein_id="AADI4516.1"
/db_xref="GI:4262223"
/translation="MRCKVYVAARTTTRTKYTKTISMGLOGLSDVSSDSITPLM
LTLALTFEHRVRSILFPSSAPVYVYVTSNLSVLADQNLNRLTFEYRSYSDNASDCIY
```

```
CLSLKLTGEEVYRKLDICRHVFNHKKQLEGMQLHNLNPNCLGRPLPHHHOCHGSDASIS
AFPLRSTASHSH"
12263..12319
/rpt_family="(CAA)n"
13681..13754
/gene="At2g01160"
/product="tRNA-Ala"
/anticodon=(pos:13715..13717,aa:Ala)
13681..13754
/gene="At2g01160"
/note="F10A8.4"
complement(join(<13980..14327,14404..14533,14615..14904,
14978..15159,15239..15480,15649..15909,15978..16109))
/gene="At2g01170"
complement(<13980..16109)
/note="F10A8.5: identical to GP.AF019637"
complement(join(13980..14327,14404..14533,14615..14904,
14978..15159,15239..15480,15649..15909,15978..16078))
/gene="At2g01170"
/codon_start=1
/product="putative amino acid permease"
/protein_id="AADI4517.1"
/db_xref="GI:4262224"
/translation="MGLGDSQSFVPMDSGVRLKEGLYKOEKRLKDLVSFNALISFS
IISVLTGITTTTYTGURFEGTIVLYGWFLAGSFIMCVGLSMAEICSSYPTSGLYYM
SMLACPRNAPPLASMMTGCKVVVCLQMAATASVDLSAQLQIVLYLLSTGNGGXYM
GSAFVIGIGHGILFTHALLNSLPIVSFISQILANMLLGLVLMILPIVSPERA
TRFEVFNFTNGGLGITSYAYIFVGLMSQYTTTGSDASHMHEEVVDKDPKRG
IISALISILFEGGVLIGISVATVDPLSLFTSNNGGVATIAETFLAPKRRSGTS
GIYCLGVVAIVAFEGGCVSSVTSNMAVAFSDGAMPSPLMHKNSREYVINAWMYS
ALISFCVMAITISGSIYARQAMYSITIGIYIAIPIILRYTLAINTVPDPSLGK
GMVGVVAIVAVYVTSVLESFVAIFPIAETLINTVPAVAGLVALTSLYMLFSARHMF
TSPISNLS"
complement(<18798..>19706)
/gene="At2g01180"
complement(<18798..>19706)
/gene="At2g01180"
/note="F10A8.6"
complement(18798..19706)
/gene="At2g01180"
/codon_start=1
/product="putative phosphatidic acid phosphatase"
/protein_id="AADI4518.1"
/db_xref="GI:4262225"
/translation="MOEIDLSVHTIKSHGVRVASKHKHMDIIVIIAIEGLNLISF
FYRYGKMDMTDLKYPFKDNTVPIMSVVYVLLPIIVPCFYLRKTCVTDLHSLIG
ILFAVLIVGTDSIKVATGCRPRPFYRRCFPRDKRELYDALGVVCHGKAAYEKCHK
SFPSGHTSMRFLTLFSLVLYSGIKRAFNGHVAKLCLVIFPLAACLVGISRVDDY
```

Query Match 29.7%: Score 81.4; DB 8; Length 84957;
Best Local Similarity 64.7%: Pred. No. 6e-06;
Matches 121; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY	88	accgccattctcgccacgggaagccgcatggaacgcccctctcaccctctc	147
DB	61589	ATCCTCTCTTCATTCATCGGACAAATCTTGGCCGCTTGAGTGTCCCTTCACCGC	61530
QY	148	ccccctcgtcactccacccaagacgatctgaagaatccgcccctcaagaagccgtcg	207
DB	61529	CTCAGCCCATGAATTTAAACACAGACGATCAACGATATGCGCGCTTACAAACCGCTGG	61470
QY	208	agtaagctgaagctcggaatgctcgtgacctaggcaccgctcaccgccaagatgctcg	267
DB	61469	AATTTCGAGTCGCGCATGTGCTTCGCTCGACAGCGCTCCACCGCAACACGCGC	61410
QY	268	tcgaccg	274
DB	61409	TCGACCG	61403

RESULT 6
A93589

LOCUS	A93589	1118 bp	DNA	Linear	PAT 22-JAN-2000
DEFINITION	Sequence 1 from Patent WO9737028.				
ACCESSION	A93589				
VERSION	A93589.1	GI:6741794			
KEYWORDS					
SOURCE					
ORGANISM	spinach.				
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
5'UTR					
CDS					
3'UTR					
BASE COUNT	261 a	264 c	258 g	335 t	
ORIGIN					
Query Match	24.8%	Score 68;	DB 6;	Length 1118;	
Best Local Similarity	67.9%	Pred. No. 0.0024;			
Matches	95;	Conservative	0;	Mismatches 45;	Indels 0;
OY	135	ctcaacccctctctccctctcgatctctctcccaagaagcattgaagaataatgcgcgc	194		
Db	163	ctctcttcttccctctccccaacacacacagctcttaactctcaagacgattctcaagaaactccgcgc	222		
OY	195	tacaagcgcgcgaagtaagtcgagtcgcgcgcatgtgtccgcgcgcctagaccgcgcaccc	254		
Db	223	gaaaaagccgcgtcgactcgcgtcaaaatccgcgcgatgttctcgctctcggaaaccggaagtaact	282		
OY	255	gccaaagcatgtccgtcgacgcgcg	274		
Db	283	gccgcatttgcctgctgctcgcg	302		
RESULT	7				
SPIRSPiR					
LOCUS					
DEFINITION	Spinacia oleracea ribose-5-phosphate isomerase mRNA, complete cds;				
ACCESSION	L43068				
VERSION	L43068				
KEYWORDS	GI:18654316				
SOURCE	spinach.				
ORGANISM	Spinacia oleracea				
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
5'UTR					
CDS					
3'UTR					
BASE COUNT	261 a	264 c	258 g	335 t	
ORIGIN					
Query Match	24.8%	Score 68;	DB 6;	Length 1118;	
Best Local Similarity	67.9%	Pred. No. 0.0024;			
Matches	95;	Conservative	0;	Mismatches 45;	Indels 0;
OY	135	ctcaacccctctctccctctcgatctctctcccaagaagcattgaagaataatgcgcgc	194		
Db	163	ctctcttcttccctctccccaacacacacagctcttaactctcaagacgattctcaagaaactccgcgc	222		
OY	195	tacaagcgcgcgaagtaagtcgagtcgcgcgcatgtgtccgcgcgcctagaccgcgcaccc	254		
Db	223	gaaaaagccgcgtcgactcgcgtcaaaatccgcgcgatgttctcgctctcggaaaccggaagtaact	282		
OY	255	gccaaagcatgtccgtcgacgcgcg	274		
Db	283	gccgcatttgcctgctgctcgcg	302		
RESULT	7				
SPIRSPiR					
LOCUS					
DEFINITION	Spinacia oleracea ribose-5-phosphate isomerase mRNA, complete cds;				
ACCESSION	L43068				
VERSION	L43068				
KEYWORDS	GI:18654316				
SOURCE	spinach.				
ORGANISM	Spinacia oleracea				
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
5'UTR					
CDS					
3'UTR					
BASE COUNT	261 a	264 c	258 g	335 t	
ORIGIN					
Query Match	24.8%	Score 68;	DB 6;	Length 1118;	
Best Local Similarity	67.9%	Pred. No. 0.0024;			
Matches	95;	Conservative	0;	Mismatches 45;	Indels 0;
OY	135	ctcaacccctctctccctctcgatctctctcccaagaagcattgaagaataatgcgcgc	194		
Db	163	ctctcttcttccctctccccaacacacacagctcttaactctcaagacgattctcaagaaactccgcgc	222		
OY	195	tacaagcgcgcgaagtaagtcgagtcgcgcgcatgtgtccgcgcgcctagaccgcgcaccc	254		
Db	223	gaaaaagccgcgtcgactcgcgtcaaaatccgcgcgatgttctcgctctcggaaaccggaagtaact	282		
OY	255	gccaaagcatgtccgtcgacgcgcg	274		
Db	283	gccgcatttgcctgctgctcgcg	302		
RESULT	7				
SPIRSPiR					
LOCUS					
DEFINITION	Spinacia oleracea ribose-5-phosphate isomerase mRNA, complete cds;				
ACCESSION	L43068				
VERSION	L43068				
KEYWORDS	GI:18654316				
SOURCE	spinach.				
ORGANISM	Spinacia oleracea				
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
5'UTR					
CDS					
3'UTR					
BASE COUNT	261 a	264 c	258 g	335 t	
ORIGIN					
Query Match	24.8%	Score 68;	DB 6;	Length 1118;	
Best Local Similarity	67.9%	Pred. No. 0.0024;			
Matches	95;	Conservative	0;	Mismatches 45;	Indels 0;
OY	135	ctcaacccctctctccctctcgatctctctcccaagaagcattgaagaataatgcgcgc	194		
Db	163	ctctcttcttccctctccccaacacacacagctcttaactctcaagacgattctcaagaaactccgcgc	222		
OY	195	tacaagcgcgcgaagtaagtcgagtcgcgcgcatgtgtccgcgcgcctagaccgcgcaccc	254		
Db	223	gaaaaagccgcgtcgactcgcgtcaaaatccgcgcgatgttctcgctctcggaaaccggaagtaact	282		
OY	255	gccaaagcatgtccgtcgacgcgcg	274		
Db	283	gccgcatttgcctgctgctcgcg	302		
RESULT	7				
SPIRSPiR					
LOCUS					
DEFINITION	Spinacia oleracea ribose-5-phosphate isomerase mRNA, complete cds;				
ACCESSION	L43068				
VERSION	L43068				
KEYWORDS	GI:18654316				
SOURCE	spinach.				
ORGANISM	Spinacia oleracea				
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
5'UTR					
CDS					
3'UTR					
BASE COUNT	261 a	264 c	258 g	335 t	
ORIGIN					
Query Match	24.8%	Score 68;	DB 6;	Length 1118;	
Best Local Similarity	67.9%	Pred. No. 0.0024;			
Matches	95;	Conservative	0;	Mismatches 45;	Indels 0;
OY	135	ctcaacccctctctccctctcgatctctctcccaagaagcattgaagaataatgcgcgc	194		
Db	163	ctctcttcttccctctccccaacacacacagctcttaactctcaagacgattctcaagaaactccgcgc	222		
OY	195	tacaagcgcgcgaagtaagtcgagtcgcgcgcatgtgtccgcgcgcctagaccgcgcaccc	254		
Db	223	gaaaaagccgcgtcgactcgcgtcaaaatccgcgcgatgttctcgctctcggaaaccggaagtaact	282		
OY	255	gccaaagcatgtccgtcgacgcgcg	274		
Db	283	gccgcatttgcctgctgctcgcg	302		
RESULT	7				
SPIRSPiR					
LOCUS					
DEFINITION	Spinacia oleracea ribose-5-phosphate isomerase mRNA, complete cds;				
ACCESSION	L43068				
VERSION	L43068				
KEYWORDS	GI:18654316				
SOURCE	spinach.				
ORGANISM	Spinacia oleracea				
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
5'UTR					
CDS					
3'UTR					
BASE COUNT	261 a	264 c	258 g	335 t	
ORIGIN					
Query Match	24.8%	Score 68;	DB 6;	Length 1118;	
Best Local Similarity	67.9%	Pred. No. 0.0024;			
Matches	95;	Conservative	0;	Mismatches 45;	Indels 0;
OY	135	ctcaacccctctctccctctcgatctctctcccaagaagcattgaagaataatgcgcgc	194		
Db	163	ctctcttcttccctctccccaacacacacagctcttaactctcaagacgattctcaagaaactccgcgc	222		
OY	195	tacaagcgcgcgaagtaagtcgagtcgcgcgcatgtgtccgcgcgcctagaccgcgcaccc	254		
Db	223	gaaaaagccgcgtcgactcgcgtcaaaatccgcgcgatgttctcgctctcggaaaccggaagtaact	282		
OY	255	gccaaagcatgtccgtcgacgcgcg	274		
Db	283	gccgcatttgcctgctgctcgcg	302		
RESULT	7				
SPIRSPiR					
LOCUS					
DEFINITION	Spinacia oleracea ribose-5-phosphate isomerase mRNA, complete cds;				
ACCESSION	L43068				
VERSION	L43068				
KEYWORDS	GI:18654316				
SOURCE	spinach.				
ORGANISM	Spinacia oleracea				
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
5'UTR					
CDS					
3'UTR					
BASE COUNT	261 a	264 c	258 g	335 t	
ORIGIN					
Query Match	24.8%	Score 68;	DB 6;	Length 1118;	
Best Local Similarity	67.9%	Pred. No. 0.0024;			
Matches	95;	Conservative	0;	Mismatches 45;	Indels 0;
OY	135	ctcaacccctctctccctctcgatctctctcccaagaagcattgaagaataatgcgcgc	194		
Db	163	ctctcttcttccctctccccaacacacacagctcttaactctcaagacgattctcaagaaactccgcgc	222		
OY	195	tacaagcgcgcgaagtaagtcgagtcgcgcgcatgtgtccgcgcgcctagaccgcgcaccc	254		
Db	223	gaaaaagccgcgtcgactcgcgtcaaaatccgcgcgatgttctcgctctcggaaaccggaagtaact	282		
OY	255	gccaaagcatgtccgtcgacgcgcg	274		
Db	283	gccgcatttgcctgctgctcgcg	302		
RESULT	7				
SPIRSPiR					
LOCUS					
DEFINITION	Spinacia oleracea ribose-5-phosphate isomerase mRNA, complete cds;				
ACCESSION	L43068				
VERSION	L43068				
KEYWORDS	GI:18654316				
SOURCE	spinach.				
ORGANISM	Spinacia oleracea				
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
5'UTR					
CDS					
3'UTR					
BASE COUNT	261 a	264 c	258 g	335 t	
ORIGIN					
Query Match	24.8%	Score 68;	DB 6;	Length 1118;	
Best Local Similarity	67.9%	Pred. No. 0.0024;			
Matches	95;	Conservative	0;	Mismatches 45;	Indels 0;
OY	135	ctcaacccctctctccctctcgatctctctcccaagaagcattgaagaataatgcgcgc	194		
Db	163	ctctcttcttccctctccccaacacacacagctcttaactctcaagacgattctcaagaaactccgcgc	222		
OY	195	tacaagcgcgcgaagtaagtcgagtcgcgcgcatgtgtccgcgcgcctagaccgcgcaccc	254		
Db	223	gaaaaagccgcgtcgactcgcgtcaaaatccgcgcgatgttctcgctctcggaaaccggaagtaact	282		
OY	255	gccaaagcatgtccgtcgacgcgcg	274		
Db	283	gccgcatttgcctgctgctcgcg	302		
RESULT	7				
SPIRSPiR					
LOCUS					
DEFINITION	Spinacia oleracea ribose-5-phosphate isomerase mRNA, complete cds;				
ACCESSION	L43068				
VERSION	L43068				
KEYWORDS	GI:18654316				
SOURCE	spinach.				
ORGANISM	Spinacia oleracea				
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
5'UTR					
CDS					
3'UTR					
BASE COUNT	261 a	264 c	258 g	335 t	
ORIGIN					
Query Match	24.8%	Score 68;	DB 6;	Length 1118;	
Best Local Similarity	67.9%	Pred. No. 0.0024;			
Matches	95;	Conservative	0;	Mismatches 45;	Indels 0;
OY	135	ctcaacccctctctccctctcgatctctctcccaagaagcattgaagaataatgcgcgc	194		
Db	163	ctctcttcttccctctccccaacacacacagctcttaactctcaagacgattctcaagaaactccgcgc	222		
OY	195	tacaagcgcgcgaagtaagtcgagtcgcgcgcatgtgtccgcgcgcctagaccgcgcaccc	254		
Db	223	gaaaaagccgcgtcgactcgcgtcaaaatccgcgcgatgttctcgctctcggaaaccggaagtaact	282		
OY	255	gccaaagcatgtccgtcgacgcgcg	274		
Db	283	gccgcatttgcctgctgctcgcg	302		
RESULT	7				
SPIRSPiR					
LOCUS					
DEFINITION	Spinacia oleracea ribose-5-phosphate isomerase mRNA, complete cds;				
ACCESSION	L43068				
VERSION	L43068				
KEYWORDS	GI:18654316				
SOURCE	spinach.				
ORGANISM	Spinacia oleracea				
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
5'UTR					
CDS					
3'UTR					
BASE COUNT	261 a	264 c	258 g	335 t	
ORIGIN					
Query Match	24.8%	Score 68;	DB 6;	Length 1118;	
Best Local Similarity	67.9%	Pred. No. 0.0024;			
Matches	95;	Conservative	0;	Mismatches 45;	Indels 0;
OY	135	ctcaacccctctctccctctcgatctctctcccaagaagcattgaagaataatgcgcgc	194		
Db	163	ctctcttcttccctctccccaacacacacagctcttaactctcaagacgattctcaagaaactccgcgc	222		
OY	195	tacaagcgcgcgaagtaagtcgagtcgcgcgcatgtgtccgcgcgcctagaccgcgcaccc	254		
Db	223	gaaaaagccgcgtcgactcgcgtcaaaatccgcgcgatgttctcgctctcggaaaccggaagtaact	282		
OY	255	gccaaagcatgtccgtcgacgcgcg	274		
Db	283	gccgcatttgcctgctgctcgcg	302		
RESULT	7				
SPIRSPiR					
LOCUS					
DEFINITION	Spinacia oleracea ribose-5-phosphate isomerase mRNA, complete cds;				

<hr/>					
MEDLINE FEATURES	96194461	Location/Qualifiers			
SOURCE		1..1118 /organism="Spinacia oleracea"			
		/db_xref="taxon:3562"			
5'/UTR		1..24 /note="putative"			
sig_peptide		25..174 /EC_number="5.3.1.6"			
CDS		25..894 /note="cytosolic precursor of chloroplast protein putative"			
		/codon_start=1			
		/product="ribose-5-phosphate isomerase"			
		/protein_id="AAL77589.1"			
		/db_xref="GI:18654317"			
		/translation="MASAAFSILPSTSFENRHATTKLNLKFLNHRNKPFPTTTIISLSPSPTPLTQDDCLKLAKEAVDVSAGVGLGTGSTAFARIGELLSACKLTNVGTPASRRTEAOASLGIPSLVDHPRIIDLADGDVPDLITAKGRGALLREKVFASSDKFIYVDPDTKLVDLGSRLAMPVEYVGFCCKYNLKQLPIFKLCGEAKLEMGSDSPYYIVDNSNYIDLPFYSIKDAEAGRISALEGVBEHGLFGMASEVIIAKTCTGSAVKTK"			
		I175...891			
mat-peptide		/note="5' end of mature subunit determined by protein sequencing putative"			
		/product="ribose-5-phosphate isomerase"			
		892..1118			
		/note="putative"			
BASE COUNT	261 a 264 c 258 g 335 t				
ORIGIN					
<hr/>					
Query Match	24.8%; Score 68; DB 8; Length 1118;				
Best Local Similarity	67.9%; Pred. No. 0.0024;				
Matches	95; Conservative	0; Mismatches	45; Indels	0; Gaps	0;
Oy	135 ctccaccctcctcccctccgtcatcctcaacccaagaagattgaagaanaatgcgccc	194			
Dd	163 CTCCTCCTCCCTGCCCAACACACTCTTAATCATGAAGAATCGAACGGCCC	222			
Oy	195 tacaaaggcgctcagtagtaagtggactcgacctgccctcagctaggaacggctcaccc	254			
Dd	223 GAATAAGCCGCtGACTCGGTCAAATCCGCCGATGGTTTCGTCCGGAACC GGAA GTACT	282			
Oy	255 gccacaagcatgcgcgttcgacctg	274			
Dd	283 GCCGCATTTGCTGTGTCCGCG	302			
<hr/>					
RESULT	8				
LOCUS	AF154658	1084 bp	mRNA	linear	PLN 15-JUN-2000
DEFINITION	Nicotiana tabaccum clone PR49 mRNA sequence.				
VERSION	AF154658				
KEYWORDS	AF154658.1 GI:8547158				
SOURCE	common tobacco.				
ORGANISM	Nicotiana tabaccum Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotsyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.				
REFERENCE AUTHORS TITLE	Sheehan,M.J., Wang,J. and Timko,M.P. Characterization of cdnas Differentially Expressed in Roots of Tobacco (Nicotiana tabaccum cv Burley 21) During the Early Stages of Alkaloid Biosynthesis Unpublished 2 (bases 1 to 1084) Sheehan,M.J., Wang,J. and Timko,M.P. Direct Submission Submitted (27-MAY-1999)	Biology,	University of Virginia,	044 Gilmer	
JOURNAL REFERENCE AUTHORS TITLE JOURNAL					

FEATURES Hall, Charlottesville, Virginia 22903, USA
Location/Qualifiers

1. 1084

/organism="Nicotiana tabacum"

/cultivar="Burley 21"

/db_xref="taxon:4097"

/clone="PR49"

/tissue_type="root"

/note="similar to ribose 5-phosphate isomerase"

BASE COUNT 286 a 245 c 250 g 302 t 1 others

ORIGIN

Query Match

Best Local Similarity 78.7%; Score 62; DB 8; Length 1084;
Pred. No. 0.031;

Matches 74; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 163 tcaccaagaagcattgagaagaatcgccgctacaaagcgcgtgagtactgtagtcgcg 222

DB 248 TATCCAGACGACACTCAAAAACCTGCGCGCATTAACCGTGTGAAATACGAAAGCG 307

QY 223 gcatgctcctcgagcaccggtccaccgc 256

DB 308 GAATGCTCTGTGTAGGCACTGTTCAACCGC 341

RESULT 9 AP003944 135296 bp DNA linear HTG 26-JUL-2001

LOCUS AP003944/c

DEFINITION Oryza sativa chromosome 6 clone OJ1126_F05, *** SEQUENCING IN

PROGRESS ***, in ordered pieces.

AP003944

VERSION AP003944.1 GI:15021914

KEYWORDS HTG; HTGS; PHASE2.

SOURCE Oryza sativa (cultivar: Nipponbare) DNA, clone: OJ1126_F05.

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzeae; Oryza.

1 (bases 1 to 135296)

Sasaki, T., Matsumoto, T. and Yamamoto, K.

Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, BAC

clone: OJ1126_F05

Published Only in Database (2001) In press

2 (bases 1 to 135296)

Sasaki, T., Matsumoto, T. and Yamamoto, K.

Direct Submission

Submitted (25-JUL-2001) Takuji Sasaki, National Institute of

Agricultural Resources, Rice Genome Research Program, Kannondai

2-1-2, Tsukuba, Ibaraki 305-8602, Japan

(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,

Tel: 81-298-38-7441, Fax: 81-298-38-7468)

The nucleotide sequence of this BAC clone was generated by

combining Monsanto and RGP-Japan sequencing data.

NOTE: It currently consists of 1 contigs. Gaps between the contigs

are represented as runs of N. The order of the pieces is believed

to be correct as given, however the sizes of the gaps between them

are based on estimates that have provided by the submitter. This

sequence will be replaced by the finished sequence as soon as it is

available and the accession number will be preserved.

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

Location/Qualifiers

1. 135296

/organism="Oryza sativa"

/cultivar="Nipponbare"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

/chromosome="6"

/clone="OJ1126_F05"

/chromosome="6"

/db_xref="taxon:4530"

Oy	61	ccttaacattaaatgagcattccctaccacccccatttcatgcggcgagaaggcgcca	120
Db	117472	CCTTCAACGGCAGCGACCGCCTGCTACGACTCCGACATCTTGCCCCCTTCGAAGCGGCCA	117531
Oy	121	tggagccggcgcttcctccaccctctcccctccgttatctactcaaccaagacatttga	180
Db	117532	TGCGCGGAGGGGTGACGCGTCAATCTCGCGCTCGTGGGCGCGACCCCAACGACTACCTCG	117591
Oy	181	agaataatgcgacctacaagcgctgcagtacgtgtagttccggcattggtctcggccta	240
Db	117592	AGGAGCGCATCGCCATCGCGCGCCCTCCACAGCGCGTCAAGCGCGCATCACCGCTCTGCT	117651
Oy	241	gcaccgctccaccgcgc 257	
Db	117652	CGCGCAGCACTTCGCGC 117668	
RESULT	11		
AP003616/c			
LOCUS			
DEFINITION	AP003616	150638 bp	DNA
ACCESSION	Oryza sativa chromosome 6 clone P0490F09, *** SEQUENCING IN PROGRESS ***, In ordered pieces.		
VERSION	AP003616	2	GI:15341241
KEYWORDS	HTG; HTGS; PHASE2.		
SOURCE	Oryza sativa (cultivar:Nipponbare) DNA, clone:P0490F09.		
ORGANISM	Oryza sativa Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartioideae; Oryzaceae; Oryza. 1 (bases 1 to 150638) Sasaki,T., Matsumoto,T. and Yamamoto,K. Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC clone:P0490F09 Published Only in Database (2001) In press 2 (bases 1 to 150638) Sasaki,T., Matsumoto,T. and Yamamoto,K. Direct Submission Submitted (10-MAY-2001) Takuji Sasaki, National Institute of Agrobiological Resources, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail:tsasakigabr.afric.go.jp, URL:http://irg.dna.afric.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468) On Aug 28, 2001 this sequence version replaced gi:14020954. NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved. * NOTE: This is a 'working draft' sequence. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved. Location/Qualifiers 1..150638 /organism="Oryza sativa" /cultivar="Nipponbare" /db_xref="taxon:4530" /chromosome="6" /clone="P0490F09"		
FEATURES			
source			
BASE COUNT	42993 a	32177 c	32647 g 42721 t 100 others
ORIGIN			
Query Match	19.9%; Score 54.6;	DB 2;	Length 150638;
Best Local Similarity	54.8%; Pred. No. 0.54;		
Matches 108;	Conservative 0;	Mismatches 89;	Indels 0; Gaps 0;
Oy	61	ccttaacattaaatgagcattccctaccacccccatttcatgcggcgagaaggcgcca	120
Db	143219	CCTTCAACGGCAGCGACCGCCTGCTACGACTCCGACATCTTGCCCCCTTCGAAGCGGCCA	143160

Oy	121	tggagcgcgacgtctccaccctccctcccttcgcgatcctcaaccaagcattga	180
Oy	181	agaaatcgccgcctcacaaaggcgtcgaatcagtggagatccgcgatggcttgacctag	240
Db	143099	AGGACGCCATGCATCGGGCCCTTCACAGCCGCTCAAGAGCCGGCATCAGCTCGTCTGCT	143040
Oy	241	gcaccggtccacgccgc 257	
Db	143039	CCGCACGCAACTTCGGC 143023	
RESULT	12		
LOCUS	AY045785	998 bp mRNA linear PLN 28-JUL-2001	
DEFINITION	Arabidopsis thaliana putative ribose 5-phosphate isomerase (F7018.28/At3g04790) mRNA, complete cds.		
ACCESSION	AY045785		
VERSION	AY045785.1	GI:15027856	
KEYWORDS	FIL_CDNA.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.		
AUTHORS	1 (bases 1 to 998) Jiang,P.X., Liu,S.X., Pham,P.K., Banh,J., Dale,J.M., Goldsmith,A.D., Toriumi,M., Yamamura,Y., Yu,G., Yu.S., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A. Full Length CDNA of gene F7018.28/At3g04790 (GI:15175179)		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 998)		
REFERENCE	Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu.S., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P., Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A. Direct Submission Submitted (10-JUL-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA		
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFLP CDNAs (RFLP CDNA : "RIKEN Arabidopsis Full-Length CDNA") : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.		
TITLE	The Salix, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RFLP CDNAs: Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Yamamura,Y., Yu.G., Yu.S., Bowser,L., Chen,H., Cheuk,R., Karlin-Neumann,G., Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Palm,C.J., Shin,P., Southwick,A., Tracy,S.E., Nguyen,M., Davis,R.W., Ecker,J.R. and Theologis,A.		
FEATURES	Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.		
SOURCE	Location/Qualifiers 1..998 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="3"		

/note="Similar to sodium proton exchanger Nhx1 GB:ABD16946
 [Arabidopsis thaliana]"
 join(8161..8531,8463..8581,9011..9108,9205..9264,
 9340..9386,9475..9669,9764..9810,9892..9958,10061..10161,
 10251..10354,10446..10584,10669..10731,10832..10915,
 11142..11489)"
 /gene="99j14.2"
 /codon_start=1
 /product="catalytic sodium proton exchanger: 8161-11489"
 /protein_id="AAG51408.1"
 /db_xref="GI:12322846"
 /translation="MTMFASLTSMKLSVSTNSDASVSLNLFVALLCACTIVGHILE
 NRMNNESTALLIGTGVVILLIRSKNQLHFLSPEDLFTYLLPPIIPNPFQVK
 KQFRNPFTTIAFGAIGTVVSCITLISGAIQFPRKDKITGTFDGDLELGAIFNAD
 VCTQLVLMDDTPLLISLIVFEGEVNDATVAFENAIQSDLTJLHNEAFOPLNGF
 YLFLFGTSHGVATGLISAVYIKLKYGRSHSTDEVLNLMIAVLSYMLAEPLASLT
 IVFEGTSLSHYTNWNTESRITTKTHAFETLSFKLETFLVGDADLDEIWRVS
 DSPGSAVNSILMGLVMGLAAPFSLTSLMSLNHOSKTSIQOQVIMAGMR
 GAVSALAAVNSITRGHTELKGNAMITSTITVCLSTFWYEGKSLPLRLYPLHOKA
 TSTSTMSLDSTPKSIHIFLIDGEOGLDSESLHSDQVPRPSLNGFLMRPTRYHY
 VMROPDPMRPVCGSGFVPGSPTRSSDLS"
 join(<12067..12114,12560..12676,12757..12879,
 12957..>13082)"
 /gene="99j14.3"
 12067..13082
 /gene="99j14.3"
 /note="identical to GB:P1829 [Arabidopsis thaliana]"
 join(12067..12114,12560..12676,12757..12879,12957..13082)"
 /gene="99j14.3"
 /codon_start=1
 /product="acyl carrier protein 1 precursor (ACP):
 12067-13082"
 /protein_id="AAG51406.1"
 /db_xref="GI:12322844"
 /translation="MATQFSASVSLQTSCLATRIQKALISNHKNTLSNPLRS
 IPSRLTSACAKQETIEKYSALYKKQLSTDPKKVVAETKFDLCLDSDLDVTEIWMG
 LEEENIMAEKAKQIATVQDAELTEULINKK"
 complement(join(<13275..13560,13812..13893,13995..14097,
 14170..14365,14560..14751,15058..>15342))
 /gene="99j14.4"
 complement(13275..15342)
 /gene="99j14.4"
 /note="similar to GB:AB61079 (contains weak similarity to
 the SAP protein (TR:E236624)) [Arabidopsis thaliana]"
 complement(join(13483..13560,13812..13893,13995..14097,
 14170..14365,14560..14751,15058..15342))
 /gene="99j14.4"
 /codon_start=1
 /product="unknown protein: 15342-13483"
 /protein_id="AAG51409.1"
 /db_xref="GI:12322847"
 /translation="MRVLSIETAESPFVYSRLSPDSTATGGLIGWVKCHFLHNTVL
 VLASITLFAVYLAAYAKKSLSKLSNRSPYIKIAYVGFMLVSLNLAMCLOAMECPG
 KEVILNLTLLTTSGLMFLFSLVLAAPFQGNVAGLMLRTFLISGLVIGDILAKA
 IYLFEGFAPLIDNNEHIIHKFKWGLVIYHLLAGIYCGMTEFMYNSKMRSEPLDGEV
 HSPYPRAPRYKITYMLALNGISLACALLDAGHFGMLTLGITISVCTHAFYPLL
 YVTRADDFQEDLNSLVYSEKKDGFEDAME"
 join(16215..16399,16873..16990,17073..17295,17473..17732)"
 /gene="99j14.5"
 16215..17732
 /gene="99j14.5"
 /note="similar to GB:CAH0547 [Caenorhabditis elegans]"
 join(16248..16399,16873..16990,17073..17295,17473..17501)"
 /gene="99j14.5"
 /codon_start=1
 /product="unknown protein: 16248-17501"
 /protein_id="AAG51412.1"
 /db_xref="GI:12322850"
 /translation="MGREVAESGVDMFMEMVAGRYANKPELARIARIGVQ
 GHQLSERTMRRPFSDLHAIKTCIDPFSEYFKQIDNLKTHNGTIVLDQNKFRW
 LSRVSDSSNSNEQDPSFGEESAAQAVSYLTFPGIIRGLSNLGIICAVASADIS
 SLTPCSFVIRKA"
 19191..19378

```

/note="Emigrat MITE-like element from T517 [17254 to
17779] 526 nt"
/gene="T9J14.6"
20266..>20949
/note="T9J14.6"
20266..>20949
/note="predicted by genefinder"
20266..>20949
/codon_start=1
/gene="T9J14.6"
/product="hypothetical protein; 20266-20949"
/protein_id="AA651415.1"
/db_xref="GI:12322853"
/translacion="MERIKADAEFTMCAQQLKGLNHLMEGANIERKSSREDLKER
ELQILSLDEQKSHAFEEKSEVGDILKLVEECTEELRSRLNLTIVKLSLIRQREL
ELKNOVOVMAELKRRYSARHVKQKRRMEDETARKKLESLTVQDIOESKOZEK
KSREVELKDKDEIEKEGKELDVKSQVKAWEKRLIQLSKLDDCCTELSPRKQDVSS
NNTYKKEIG"
<22764..>22621
/gene="T9J14.7"
22764..26261
/gene="T9J14.7"
/note="similar to GB:AA61072 [Arabidopsis thaliana]"
22764..26261
/gene="T9J14.7"
/codon_start=1
/product="hypothetical protein, contains DnaJ motif;
prokaryotic heat shock protein motif; 22764-26261"
/protein_id="AA651418.1"
/db_xref="GI:12322856"
/translacion="MEAGDFVGAKHFVTKAQRLEPNIENIVOMMTICDVHSAIKIK
GLDQVGLVQPYPADADITTKQYRKLAHLHPDKNFAGAEAFKLVGANLISQ
IKRQYDNRVSRHSMFANRHNHNVSGRHCAATNNAENIGVFTFTRCHCQCYK
LREWMNTHMOSSQCSKFVACKMKDCDEVPSSSTAGRKEPDQVMSTSHQNSTAAE

```


TITLE
JOURNAL Arabidopsis thaliana chromosome III BAC F7018 genomic sequence
REFERENCE
2 (bases 1 to 95310)
AUTHORS Lin, X. and Kaul, S.
TITLE Direct Submission
JOURNAL Submitted (06-OCT-1999) The Institute for Genomic Research, 9712
REFERENCE Medical Center Dr., Rockville, MD 20850, USA, xlin@tigr.org
AUTHORS 3 (bases 1 to 95310)
TITLE Lin, X.
JOURNAL Direct Submission
COMMENT Submitted (24-JAN-2001) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Jan 24, 2001 this sequence version replaced gi:12280830.
Addresses all correspondence to:
Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr.
Rockville, MD 20850, USA
e-mail: xlin@tigr.org
BAC clone F7018 is from Arabidopsis chromosome III and is near the
molecular marker ml172.
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.

Genes were identified by a combination of three methods: Gene
prediction programs including GRAFT (available by anonymous ftp
from arthur.epm.cornl.gov), GeneFinder (Phil Green, University of
Washington), Genscan (Chris Burge,
http://genome.csf.cmu.edu/~chris/burges/genscan.html), and NetPlantene
(http://www.cds.dtu.dk/netgene/cbsnetgene.html), searches of the
complete sequence against a peptide database and the Arabidopsis
EST database at TIGR (http://www.tigr.org/cdb/at/est.html).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as 'unknown' proteins.
Genes without protein or EST similarity, that are predicted by more
than two gene prediction programs over most of their length are
annotated as 'hypothetical' proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are
identified by RepeatMasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html). Regions of
genomic sequence that are not annotated as genes but have predicted
exons by GRAFT are annotated as misc features.
Location/Qualifiers

source
1. 95310
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="III"
/map="ml172"
/clone="F7018"
join(<1. 691,1263. >1321)
/gene="F7018.1"
<1. >1321
/gene="F7018.1"
/note="Similar to calmodulin-domain protein kinase (CDPK
isoform 9) GB:AAB03242 [Arabidopsis thaliana], similar to
hypothetical protein (P22013.13) GB:AAC14039 [Arabidopsis
thaliana]; Pfam HMM hit: Eukaryotic protein kinase domain"
join(<1. 691,1263. 1321)
/gene="F7018.1"
/codon_start=1
/product="putative calmodulin-domain protein kinase, 5'
partial"
/protein_id="AAF04907.1"
/db_xref="GI:6175181"
/translation="EFFACKTIDKRYLDALDRECIETEPRIAMLPHPNIRIFDL
YETEDSLAIVMELVDPPIYDRLISAGRLSESSASAYAKOILSALAHCHGCDVYHR
DVKPDNVLIVLWGGVYKLCDFGSAMVLDGEGTAVGVGTPYVAAPEVWGRKXDEVDT
WSAGVVIYTMALGEPFNGETAFEDFESTLRGNLPPPKKFSVSSEAKDLRKMKIC
DVSRRSADDAHLVETLNTQVIRLDSNVY"
repeat_region
complement(700. 730)

misc_feature
1497. 1548
/rpt_family="POLY_A"
/note="exon predicted by xgrail, quality good"
misc_feature
1658. 1742
/note="exon predicted by xgrail, quality excellent"
misc_feature
complement(1692. 1764)
/note="exon predicted by xgrail, quality
marginal_shadowexon"
repeat_region
complement(1811. 1888)
/rpt_family="(TA)n"
repeat_region
complement(1812. 1908)
/rpt_family="(CARA)n"
1910. 2023
/note="exon predicted by xgrail, quality excellent"
misc_feature
complement(1965. 2063)
/note="exon predicted by xgrail, quality
marginal_shadowexon"
misc_feature
2787. 2833
/note="exon predicted by xgrail, quality good"
complement(3313. 3418)
/note="exon predicted by xgrail, quality
excellent_shadowexon"
misc_feature
3319. 3403
/note="exon predicted by xgrail, quality good"
3518. 3631
/note="exon predicted by xgrail, quality excellent"
complement(3573. 3723)
/note="exon predicted by xgrail, quality
marginal_shadowexon"
3957. 4004
/note="exon predicted by xgrail, quality marginal"
<4314. 5760
/gene="F7018.2"
<4314. 5760
/gene="F7018.2"
/note="predicted by genscan, multiple est matches"
4314. 5663
/gene="F7018.2"
/note="unknown protein"
/codon_start=1
/protein_id="AAF04886.1"
/db_xref="GI:6175160"
/translation="MFSLSLSISPFQSTHGLFTNPTRPVNPDPRTVSFTYFASM
IPKRSANMIPKPNPPOOLYOPPEPSPPIPOFRLSDSGKTEIAGHMAIFEA
PLISLYTDGFTPTTELTGISTEONRLTGVQVDSITQSIHEPELISAEPTGA
ELYLEIRLSTTORVAATFTIDRIDSKGQADAKAKDIPNRGQVGMDEVDYNLP
GDCLSEIYVROSREKNKNSDQRTSMLDALGVASEKAKNLTLEAMEIAGDFVY
KKKEEVAKIRLIPVRLKFGVAKETSVVVLPCVKAEGEKRLLEAMEIAGDFVY
VEAKGMRWVVLPSMNPVAAIGGVAVSFRRDRKVLPMGKDEPLLVVADRNVY
EADGGLVVAENGKLEKESGDLKAREVESLGVVLVVRPRDDDDQWQTSHQND"
5066. 5160
/rpt_family="(GAA)n"
complement(join(5740. 6123,6229. 6587,6767. 6856,
6946. 7229,7314. 7391,7514. >7691))
/gene="F7018.3"
complement(5740. >7691)
/gene="F7018.3"
/note="predicted by genscan, multiple est matches"
complement(join(5859. 6123,6229. 6587,6767. 6856,
6946. 7229,7314. 7391,7514. 7691))
/gene="F7018.3"
/note="unknown protein"
/codon_start=1
/product="putative calmodulin-domain protein kinase, 5'
partial"
/protein_id="AAF04887.1"
/db_xref="GI:6175181"
/translation="MAENLFGPDPRESSQQQELPISLPDDESKITSSPAPILVLS
ALKRSPEBSAPNLSAPVLKSLAKRSKPESTPEVPEEAPKRRKLOKSTDSSE
QVTAWKITSHIKNPSKFSKASKLAIRLQAGSVPESTSYIALIEAAMSSTPCT
KRSADRVAAHLESAQDAVACDLKRSQNLITITFAVVAANDLFTDSDSPFKATRI
DEASIDLVSTEDDYDEAAALEAAVVKDGDQGTQOVAAEASAGDNAAVSDPRL
DAWIPSSKNGKTKIKRTNEDDAENKRFSLSKREALITLTETLAARYKPKQCTV
IDILVKAFFENVSFTSQORAVETKIMASVREHLLRKKGKSVTGKLDVTAFESLDQK
YANEKMSIRSSVAGSGERRAQDWLG"

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 03:38:57 ; Search time 6582.52 Seconds
(without alignments)
561.816 Million cell updates/sec

Title: US-09-300-482-311

Perfect score: 274
Sequence: 1 cttaacatctcttctccacc.....gccaaagcatgcgcgcgacgcg 274

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlma:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estc2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vtl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189.6	69.2	569	10	BM093169
2	141.6	51.7	411	9	AV412358
3	133.6	48.8	520	10	BF645504
4	133.6	48.8	562	10	BF520202
5	133.6	48.8	563	9	AM559864
6	133.6	48.8	668	9	AM559587
7	133.6	48.8	679	10	BG647838
8	133.6	48.8	704	10	BG584469
9	133.6	48.8	808	10	BG582993
10	126.4	46.1	428	9	AI900328
11	112	40.9	275	9	AV424584
12	107.2	39.1	722	10	BG585240
13	102.8	37.5	636	10	BI309133
14	99	36.1	265	9	AV412159
15	96.8	35.3	266	9	AV412011
16	94.2	34.4	457	10	BI139102

18	88	32.1	309	9	AV416699
19	88	32.1	366	9	AV420477
20	88	32.1	374	9	AV414597
21	88	32.1	379	9	AV421081
22	88	32.1	382	9	AV415702
23	88	32.1	385	9	AV424000
24	88	32.1	389	9	AV407718
25	88	32.1	395	9	AV413372
26	88	32.1	405	9	AV422025
27	88	32.1	422	9	AV425897
28	88	32.1	425	9	AV426595
29	88	32.1	426	9	AV407869
30	88	32.1	428	9	AV418494
31	88	32.1	429	9	AV411127
32	88	32.1	466	10	BI417490
33	88	32.1	471	9	AV422350
34	88	32.1	554	10	BI420175
35	88	32.1	571	10	BI418851
36	88	32.1	573	9	AM720035
37	87.6	32.0	481	10	BF067532
38	87.6	32.0	501	10	BI498790
39	87.6	32.0	502	10	BG363532
40	87.6	32.0	595	10	BF067604
41	87.6	32.0	604	10	BG042907
42	87.6	32.0	611	9	BE191266
43	86.6	31.6	322	9	AV428885
44	86	31.4	368	9	AI460598
45	86	31.4	395	10	BE609370

ALIGNMENTS

RESULT 1
LOCUS BM093169
DEFINITION saj05h04.y1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl065-9584.5' similar to SW:RPIA_ARATH G92U03 PROBABLE
RIBOSE 5-PHOSPHATE ISOMERASE ; mRNA sequence.

ACCESSION BM093169
VERSION BM093169
KEYWORDS
SOURCE
ORGANISM
soybean.

REFERENCE 1 (bases 1 to 569)
Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
'A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
'Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk
'R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
'R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 421.
Location/Qualifiers
1..569
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-9584"

TITLE
JOURNAL
COMMENT

FEATURES
source


```

/dey_stages="cotyledons and primary leaves harvested 5 and
8 days after inoculation with Colletotrichum trifolii"
/lab_host="E. coli strain XL04R"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA from
cotyledons and primary leaves harvested 5 and 8 days after
inoculation with Colletotrichum trifolii. The cDNA was
directionally ligated into the Uni-ZAP XR vector from

```

```

/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pDSIR-24p14"
/clone_1lb="DSIR"
/tissue_type="roots infected with Phytophthora
medicaginis"
/dev_stage="roots harvested at 10 days post inoculation
with Phytophthora medicaginis"

```


/clone="pGPOD-10D22"
/clone_lib="GPOD"
/tissue_type="Immature pod walls"
/dev_stage="Immature pods, ranging in age from 15 to 30 days after pollination"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Immature pods, ranging in age from 15 to 30 days after pollination, were collected from greenhouse-grown plants. At harvest, seeds were removed from pods and isolated pod walls were collected and immediately frozen in liquid nitrogen. Pod walls were pooled for mRNA extraction. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL0LR cells."

BASE COUNT 157 a 135 c 149 g 195 t

ORIGIN

Query Match 37.5%; Score 102.8; DB 10; Length 636;
Best Local Similarity 84.1%; Pred. No. 2.5e-11;
Matches 116; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

OY 137 caaccctctccctccctccgtcattccaccacaagattgaagaataatgcgcctca 196
DB 24 cccattctctccgttcgttcattccatccacagacattgaagaaatgcgcctca 83

OY 197 caagcgctcgagtagtgagtcgagcatgtctctcgacgacgacgctccacgc 256
DB 84 CAAAGCCGTCGAATAGTCGATCCGGATGTTAGAGCTCGAAGCCGTTCAACCGC 143

OY 257 caagcatgcgctgcacgcg 274
DB 144 CAAACACGCGCTCGCCG 161

RESULT 14
AV412159 265 bp mRNA linear EST 23-MAY-2000
LOCUS AV412159 Lotus japonicus young plants (two-week old) Lotus
DEFINITION AV412159 Lotus japonicus young plants (two-week old) Lotus
ACCESSION AV412159
VERSION AV412159.1 GI:7741323
KEYWORDS EST.
SOURCE Lotus japonicus.
ORGANISM Lotus japonicus.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.

REFERENCE 1 (bases 1 to 265)
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE Generation of 7137 non-redundant expressed sequence tags from a
JOURNAL legume, Lotus japonicus
MEDLINE DNA Res. 7 (2), 127-130 (2000)
COMMENT Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakam@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1..265
/organism="Lotus japonicus"
/db_xref="taxon:34305"
/clone_lib="MM21610_r"
/dev_stage="young plants (two-week old)"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI; isolate= Miyakojima MG-20"

BASE COUNT 54 a 92 c 31 g 88 t

ORIGIN

Query Match 36.1%; Score 99; DB 9; Length 265;
Best Local Similarity 78.7%; Pred. No. 1.3e-10;
Matches 133; Conservative 0; Mismatches 30; Indels 6; Gaps 1;

OY 57 tcaaccttaacatnaacatgccaatccaccacattcatcgccacgaagaagcc 116
DB 103 TCTCTCTCGATTTCATATGCTATTCCTATCCCATTTTCATCTCTCCGAGAAAGCA 162

OY 117 gccatgagcgccgctctccaccctcccccctccgcgtcatctccaccacgaagat 176
DB 163 GCCATGGAAGACGAGGCTTCTCGCACCAT-----CTTCATCATTTCTCACCAAGAGAT 216

OY 177 ttgaagaataatgcgcgctcacaagcgctgagtagtgagtcgcgca 225
DB 217 TTGAAGAAATCGCCGCTTACAAAGCCGTGAGTATGTCGATTCGGCA 265

RESULT 15
AV412011 266 bp mRNA linear EST 23-MAY-2000
LOCUS AV412011 Lotus japonicus young plants (two-week old) Lotus
DEFINITION AV412011 Lotus japonicus young plants (two-week old) Lotus
ACCESSION AV412011
VERSION AV412011.1 GI:7741175
KEYWORDS EST.
SOURCE Lotus japonicus.
ORGANISM Lotus japonicus.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.

REFERENCE 1 (bases 1 to 266)
AUTHORS Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE Generation of 7137 non-redundant expressed sequence tags from a
JOURNAL legume, Lotus japonicus
MEDLINE DNA Res. 7 (2), 127-130 (2000)
COMMENT Contact: Yasukazu Nakamura
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: ynakam@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1..266
/organism="Lotus japonicus"
/db_xref="taxon:34305"
/clone_lib="MM214a02_r"
/dev_stage="young plants (two-week old)"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI; isolate= Miyakojima MG-20"

BASE COUNT 54 a 94 c 30 g 88 t

ORIGIN

Query Match 36.1%; Score 99; DB 9; Length 266;
Best Local Similarity 78.7%; Pred. No. 1.3e-10;
Matches 133; Conservative 0; Mismatches 30; Indels 6; Gaps 1;

OY 57 tcaaccttaacatnaacatgccaatccaccacattcatcgccacgaagaagcc 116
DB 104 TCTCTCTCGATTTCATATGCTATTCCTATCCCATTTTCATCTCTCCGAGAAAGCA 163

OY 117 gccatgagcgccgctctccaccctccctccctccgtcattcatccaccacgaagat 176
DB 164 GCCATGGAAGACGAGGCTTCTCGCACCAT-----CTTCATCATTTCTCACCAAGAGAT 217

OY 177 ttgaagaataatgcgcgctcacaagcgctgagtagtgagtcgcgca 225
DB 218 TTGAAGAAATCGCCGCTTACAAAGCCGTGAGTATGTCGATTCGGCA 266

Wed Jul 3 12:01:15 2002

us-09-300-482-311.rst

Page 9

Search completed: July 3, 2002, 03:39:00
Job time: 43320 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 07:57:43 ; Search time 796.36 Seconds

(without alignments)
590.731 Million cell updates/sec

Title: US-09-300-482-311

Perfect score: 274
1 cttaacatcccttcaccc.....gccaaagcatgcgcgcacgc 274

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

```
1: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	81.4	29.7	942	21	AAC37420	Arabidopsis thalia
2	81.4	29.7	1429	21	AAC47590	Arabidopsis thalia
3	81.4	29.7	1369	21	AAC47074	Arabidopsis thalia
4	68	24.8	1118	18	AAC87859	DNA encoding ribos
5	54	19.7	966	21	AAC45894	Arabidopsis thalia
6	54	19.7	967	21	AAAT8440	Plant SDF polynuc
7	49.2	18.0	454	22	AAH88764	Ribose-5-Page-ism
8	44.6	16.3	77536	21	AAH14651	Nucleotide sequenc
9	43.4	15.8	1233	22	AAH44089	Oryza sativa perox

10	43	15.7	701	18	AAAT94500	GAL4 DNA binding d
11	43	15.7	38734	20	AAAT32020	Human MERT1 relate
12	43	15.7	38734	22	AAC90077	AL021529 cDNA clon
13	42.6	15.5	1371	21	AAAT87214	DNA encoding synth
14	42.6	15.5	1371	22	AAAT54486	Botulin toxin hea
15	42	15.3	771	20	AAAT12317	Neisseria meningit
16	41.8	15.3	833	15	AAAT64203	snab gene encoding
17	41.8	15.3	882	19	AAAT64512	M. tuberculosis im
18	41.8	15.3	882	19	AAAT44403	M. tuberculosis im
19	41.8	15.3	882	20	AAAT19313	M. tuberculosis an
20	41.8	15.3	882	20	AAAT19101	M. tuberculosis re
21	41.8	15.3	5392	15	AAAT64201	Sequence compisrin
22	41.8	15.3	4411529	22	AAAT99682	Mycobacterium tube
23	41.6	15.2	290	18	AAAT91500	Mycobacterium tube
24	41.6	15.2	290	18	AAAT91436	Mycobacterium tube
25	41.6	15.2	290	19	AAAT64483	M. tuberculosis im
26	41.6	15.2	290	19	AAAT4375	Mycobacterium tube
27	41.6	15.2	290	20	AAAT19285	M. tuberculosis an
28	41.6	15.2	290	20	AAAT19073	M. tuberculosis re
29	41.6	15.2	1929	19	AAAT57472	Sorghum bicolor (L
30	41.4	15.1	490	21	AAAT9485	Eucalyptus grandis
31	41.4	15.1	4403765	22	AAAT9683	Mycobacterium tube
32	41	15.0	3632	17	AAAT06978	T. thermophilus ga
33	41	15.0	3632	20	AAAT21373	Thermus thermophil
34	41	15.0	3849	22	AAAT5795	S. chrysomallus ac
35	40.8	14.9	985	19	AAAT64548	M. tuberculosis im
36	40.8	14.9	985	19	AAAT44439	Mycobacterium tube
37	40.8	14.9	985	20	AAAT19349	M. tuberculosis an
38	40.8	14.9	985	20	AAAT19137	M. tuberculosis re
39	40.8	14.9	27541	22	AAAT17185	Streptomyces nous
40	40.8	14.9	125401	22	AAAT17186	Streptomyces nous
41	40.4	14.7	1236	16	AAAT5060	Waite T52 cDNA nuc
42	40.4	14.7	1248	21	AAAT87286	S. venezuelae deso
43	40.4	14.7	1600	22	AAAT4793	Cysteine protease
44	40.4	14.7	5970	21	AAAT5635	Nucleotide sequenc
45	40.4	14.7	5970	21	AAAT56003	Contig 002 from co

ALIGNMENTS

```
RESULT 1
AAC37420
ID AAC37420 standard; DNA; 942 BP.
XX
AC AAC37420;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 17312.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 09-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
```


PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134769.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140961.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145291.
PR 27-JUL-1999; 99US-0145291.
PR 27-JUL-1999; 99US-0145291.
PR 28-JUL-1999; 99US-0145291.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144634.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144886.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150586.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152353.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158393.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.

PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 29.7%; Score 81.4; DB 21; Length 1429;
Best Local Similarity 64.7%; Pred. No. 3.7e-11;
Matches 121; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 88 accccatttcacgcacgcagaaagcgcgcacatggacgcgcctctccaccctct 147
Db 368 atctctcttcattacacgcagacaatcttgcgcgcttggatgtgctcttcacgcg 427
QY 148 cccctccgcacatccctaccccaagacgatttgaagaatacgcgcctcaagccgtcg 207
Db 428 ctacgccacgaatttaacacagacgagctcaacgctacgcgcttacaagccgttg 487
QY 208 agtacgtgagtcgcgcacatgctcctgcgcctagcacgcgcctccacgcgaagatcgcg 267
Db 488 aattcgtcgagtcgcgcacatgctcctgcgcctagcacgcgcctccacgcgaagatcgcg 547
QY 268 tcgaccg 274
Db 548 tcgaccg 554

RESULT 3

ID AAC47074 standard; DNA; 1569 BP.

XX AAC47074;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 52467.

KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132853.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139839.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145291.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149930.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.


```
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 29.7%; Score 81.4; DB 21; Length 1569;
Best Local Similarity 64.7%; Pred. No. 3.8e-11;
Matches 121; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
```

```
QY 88 accccattcatcgccagaaagccgcgcatgagcgcgcctctccaccctctc 147
   ||| ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 508 atccctcttcatctacgacaatcttctgcgccttgaagtgtcccttcacgc 567
QY 148 cccctcgtcatctcctcccaagacgaattgaagaatacgccctacaaagcgctc 207
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 568 ctcaagccatgaatttaacacagacgagctcaacgcatcgccgttaacaagcgctg 627
QY 208 agtaagtgaagtcgcagctgtctcgcctcagcagcgcgttcacccgcaagatgcg 267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 628 aatcgctgaagtcgcagctgttctcggtctcgcgaaccgcgtccacgcgcaaacgcg 687
QY 268 tcgacgcg 274
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 688 tcgacgcg 694
```

```
RESULT 4
AAT87859
ID AAT87859 standard; cDNA; 1118 BP.
AC AAT87859;
XX
XX 21-APR-1998 (first entry)
DE DNA encoding ribose-5-phosphate isomerase.
XX Ribose-5-phosphate isomerase; screen; inhibitor; herbicidal agent;
XX ss.
XX Splinacla oleracea.
XX
XX Key Location/Qualifiers
FH CDS 25..894
FT /*lag= a
FT /product= ribose-5-phosphate_isomerase
XX
XX W09737028-A2.
```

```
XX
PD 09-OCT-1997.
XX
XX 26-MAR-1997; 97WO-EP01539.
PF
XX 29-MAR-1996; 96DE-4012772.
PR
XX (BADI ) BASF AG.
XX
XX Kellermann J, Lerchl J, Schmidt R, Schnarrenberger C;
PI William M;
XX WPI: 1997-503113/46.
DR P-PDSB: AAM27505.
XX
XX Spinach ribose-5-phosphate isomerase - useful to screen for
PT inhibitors, i.e. herbicidal agents
XX
XX Claim 5; Figure 2; 22pp; German.
PS
XX The present sequence represents DNA encoding a novel spinach ribose-5-
CC phosphate isomerase. The invention relates to a protein with ribose-5-
CC phosphate isomerase activity, comprising at least 100 residues of a 289
CC residue spinach amino sequence, preferably residues 50-239. Novel
CC phosphate isomerase inhibitors can be identified by incubating potential
CC inhibitors with the enzyme and determining the ribose-5-phosphate
CC isomerase activity. These inhibitors can be used as herbicidal agents.
XX
SQ Sequence 1118 BP; 261 A; 264 C; 258 G; 335 T; 0 other;
```

```
Query Match 24.8%; Score 68; DB 18; Length 1118;
Best Local Similarity 67.9%; Pred. No. 7.9e-08;
Matches 95; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 135 ctccaccctctctccctcgtcatctctcaccacgaagatctgaagaatcgccgc 194
   ||| ||| ||||| ||| ||| ||||| ||| ||| ||||| ||||| ||||| |||||
Db 163 ctctctctctccctccccaacacagctcttaactcaagacgatctcaagaactcgccgc 222
QY 195 tacaagcgcgtgaagtagtgagtcgcgcatgtctctcgctcgaagcgcgctccacc 254
   ||| ||| ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 223 gaataaacgcgtcgaactcgcaaatccgcatgttctcgtctcgaaccggaagtaact 282
QY 255 gcccaagatgcgcgtcgaccg 274
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 gccgcattgtgtctcgcg 302
```

```
RESULT 5
AAC45894
ID AAC45894 standard; DNA; 966 BP.
XX
XX AAC45894;
XX
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 48154.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
```


PR	23-MAR-1999	9905-0125788
PR	25-MAR-1999	9905-0126264
PR	29-MAR-1999	9905-0126785
PR	01-APR-1999	9905-0127462
PR	06-APR-1999	9905-0128234
PR	08-APR-1999	9905-0128945
PR	16-APR-1999	9905-0129814
PR	19-APR-1999	9905-0130077
PR	21-APR-1999	9905-0130449
PR	23-APR-1999	9905-0130510
PR	28-APR-1999	9905-0130891
PR	30-APR-1999	9905-0132048
PR	04-MAY-1999	9905-0132484
PR	05-MAY-1999	9905-0132485
PR	06-MAY-1999	9905-0132486
PR	06-MAY-1999	9905-0132487
PR	07-MAY-1999	9905-0132683
PR	11-MAY-1999	9905-0134256
PR	14-MAY-1999	9905-0134218
PR	14-MAY-1999	9905-0134219
PR	14-MAY-1999	9905-0134321
PR	14-MAY-1999	9905-0134370
PR	18-MAY-1999	9905-0134768
PR	19-MAY-1999	9905-0134941
PR	20-MAY-1999	9905-0135124
PR	21-MAY-1999	9905-0135353
PR	24-MAY-1999	9905-0135629
PR	27-MAY-1999	9905-0136021
PR	28-MAY-1999	9905-0136392
PR	01-JUN-1999	9905-0136782
PR	03-JUN-1999	9905-0137222
PR	04-JUN-1999	9905-0137528
PR	07-JUN-1999	9905-0137502
PR	08-JUN-1999	9905-0137724
PR	10-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138540
PR	14-JUN-1999	9905-0138847
PR	16-JUN-1999	9905-0139119
PR	16-JUN-1999	9905-0139452
PR	16-JUN-1999	9905-0139493
PR	17-JUN-1999	9905-0139692
PR	18-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139456
PR	18-JUN-1999	9905-0139457
PR	18-JUN-1999	9905-0139458
PR	18-JUN-1999	9905-0139459
PR	18-JUN-1999	9905-0139460
PR	18-JUN-1999	9905-0139461
PR	18-JUN-1999	9905-0139462
PR	18-JUN-1999	9905-0139463
PR	18-JUN-1999	9905-0139750
PR	18-JUN-1999	9905-0139763
PR	21-JUN-1999	9905-0139817
PR	22-JUN-1999	9905-0139859
PR	23-JUN-1999	9905-0140353
PR	23-JUN-1999	9905-0140354
PR	24-JUN-1999	9905-0140695
PR	29-JUN-1999	9905-0140823
PR	30-JUN-1999	9905-0140991
PR	01-JUL-1999	9905-0141287
PR	01-JUL-1999	9905-0141854
PR	02-JUL-1999	9905-0142554
PR	06-JUL-1999	9905-0142055
PR	08-JUL-1999	9905-0142390
PR	09-JUL-1999	9905-0142603
PR	12-JUL-1999	9905-0142920
PR	13-JUL-1999	9905-0142977
PR	14-JUL-1999	9905-0143342
PR	15-JUL-1999	9905-0143624
PR	15-JUL-1999	9905-0144005

PR	16-JUL-1999	9905-0144085
PR	16-JUL-1999	9905-0144086
PR	19-JUL-1999	9905-0144325
PR	19-JUL-1999	9905-0144331
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144333
PR	19-JUL-1999	9905-0144334
PR	19-JUL-1999	9905-0144335
PR	20-JUL-1999	9905-0144352
PR	20-JUL-1999	9905-0144632
PR	20-JUL-1999	9905-0144684
PR	21-JUL-1999	9905-0144814
PR	21-JUL-1999	9905-0145086
PR	21-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145085
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145089
PR	22-JUL-1999	9905-0145192
PR	23-JUL-1999	9905-0145145
PR	23-JUL-1999	9905-0145218
PR	23-JUL-1999	9905-0145524
PR	26-JUL-1999	9905-0145576
PR	27-JUL-1999	9905-0145913
PR	27-JUL-1999	9905-0145918
PR	27-JUL-1999	9905-0145919
PR	28-JUL-1999	9905-0145951
PR	02-AUG-1999	9905-0146586
PR	02-AUG-1999	9905-0146588
PR	02-AUG-1999	9905-0146589
PR	03-AUG-1999	9905-0147038
PR	04-AUG-1999	9905-0147204
PR	04-AUG-1999	9905-0147302
PR	05-AUG-1999	9905-0147192
PR	05-AUG-1999	9905-0147260
PR	06-AUG-1999	9905-0147263
PR	06-AUG-1999	9905-0147416
PR	09-AUG-1999	9905-0147493
PR	09-AUG-1999	9905-0147935
PR	10-AUG-1999	9905-0148171
PR	11-AUG-1999	9905-0148341
PR	11-AUG-1999	9905-0148344
PR	13-AUG-1999	9905-0148654
PR	13-AUG-1999	9905-0148685
PR	15-AUG-1999	9905-0149368
PR	17-AUG-1999	9905-0149175
PR	18-AUG-1999	9905-0149426
PR	20-AUG-1999	9905-0149723
PR	20-AUG-1999	9905-0149929
PR	23-AUG-1999	9905-0149902
PR	23-AUG-1999	9905-0149930
PR	25-AUG-1999	9905-0150566
PR	26-AUG-1999	9905-0150864
PR	27-AUG-1999	9905-0151065
PR	27-AUG-1999	9905-0151066
PR	27-AUG-1999	9905-0151080
PR	30-AUG-1999	9905-0151303
PR	31-AUG-1999	9905-0151438
PR	01-SEP-1999	9905-0151930
PR	07-SEP-1999	9905-0152363
PR	10-SEP-1999	9905-0153070
PR	13-SEP-1999	9905-0153718
PR	15-SEP-1999	9905-0154058
PR	16-SEP-1999	9905-0154039
PR	22-SEP-1999	9905-01547179
PR	22-SEP-1999	9905-01551319
PR	24-SEP-1999	9905-01554869
PR	28-SEP-1999	9905-0156458
PR	28-SEP-1999	9905-0156596
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157753
PR	06-OCT-1999	9905-0157665


```
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159658.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 23-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 19.7%; Score 54; DB 21; Length 966;
Best Local Similarity 52.7%; Pred. No. 0.00024;
Matches 117; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
```

```
QY 52 tagctcaacctttaacataatgacatccctaccccatcttcacgacagaga 111
DB 58 tcgtttccctcttcacattacgtagcactcttcatttccttaagcagcactgct 117
QY 112 aagcgcatgagcgcgccctccacacccctctccctccgtcattcctaccaga 171
DB 118 ctctctcctgaacctccgcttcatctccgtcaaggcttaacccgcttcacag 177
QY 172 acgatttgaagaataatcgcgccctacaaagccgtcgaagtaagtcgagcgcgcatgctc 231
DB 178 acgatttgaagaagctcgcgcggaagaagctgtgaggaatataacttggatgttc 237
QY 232 tcggcctagacacggcgccacccacgcaagcagtcgcgtgacc 273
DB 238 tagggctcggaacgagatccacgacgagcttgcgtgtgacc 279
```

```
RESULT 6
AAA78440
ID AAA78440 standard; DNA; 967 BP.
AC AAA78440;
XX
XX 27-NOV-2000 (first entry)
DE Plant SDF polynucleotide sequence SPQ List 1 NO:25.
XX
XX Plant: corn; Arabidopsis thaliana; sequence-determined DNA fragment;
KW SDF: genetic mapping; identification; promoter; structural gene; UTR;
KW untranslated region; expression control; ds.
XX Plant.
OS
XX
XX WO2000040695-A2.
XX
XX 13-JUL-2000.
XX
```

```
PF 07-JAN-2000; 2000WO-US00466.
XX
XX 08-JAN-1999; 99US-0115293.
XX
XX (CERE-) CERES INC.
XX
XX Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L;
PI WPI, 2000-465970/40.
XX
XX New corn plant and Arabidopsis thaliana sequence-determined DNA
PT fragments, useful for expressing gene products and for controlling
PT expression of a target gene.
XX
XX Claim 1: Page 337; 673pp; English.
XX
XX The present invention describes polynucleotides, such as complete cDNA
CC sequences and/or sequences of genomic DNA encompassing complete genes,
CC portions of genes, and/or intergenic regions, collectively referred to
CC as sequence-determined DNA fragments (SDFs), from corn plants and
CC Arabidopsis thaliana. The SDFs are promoters, structural genes,
CC untranslated regions (UTRs), or 3' termination sequences. They can be
CC used for expressing a gene product and controlling expression of a
CC target gene, either as a promoter, a structural gene, an UTR or as a
CC 3' termination sequence. They are also useful as tools for genetic
CC mapping, and identification of a particular individual plant or for
CC clustering a group of plants with a common trait. AAA78433 to AAA78630
CC and AAB24605 to AAB25099 represent the specifically claimed
CC polynucleotide sequences and polypeptides encoded by them given in the
CC present invention.
XX
XX Sequence 967 BP; 232 A; 201 C; 229 G; 305 T; 0 other;
XX
XX
```

```
Query Match 19.7%; Score 54; DB 21; Length 967;
Best Local Similarity 52.7%; Pred. No. 0.00024;
Matches 117; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 52 tagctcaacctttaacataatgacatccctaccccatcttcacgacagaga 111
DB 58 tcgtttccctcttcacattacgtagcactcttcatttccttaagcagcactgct 117
QY 112 aagcgcatgagcgcgccctccacacccctctccctccgtcattcctaccaga 171
DB 118 ctctctcctgaacctccgcttcatctccgtcaaggcttaacccgcttcacag 177
QY 172 acgatttgaagaataatcgcgccctacaaagccgtcgaagtaagtcgagcgcgcatgctc 231
DB 178 acgatttgaagaagctcgcgcggaagaagctgtggaagcaatataacttggatgttc 237
QY 232 tcggcctagacacggcgccacccacgcaagcagtcgcgtgacc 273
DB 238 tagggctcggaacgagatccacgacgagcttgcgtgtgacc 279
```

```
RESULT 7
AAH8764
ID AAH8764 standard; CDNA; 454 BP.
AC AAH8764;
XX
XX 28-SEP-2001 (first entry)
DE Ribose-5-Phosphate isomerase cDNA sequence #57.
XX
XX Moss; carbohydrate metabolism related protein; GMP; sugar; cofactor;
KW fine chemical production; carbohydrate; polysaccharide; ss.
XX
XX Physcomitrella patens.
OS
XX
XX WO200144476-A2.
XX
XX
```



```

FT      /note= "KR5"
FT      complement (20241..21420)
FT      /tag= V
FT      /note= "DH5"
FT      complement (21420..22653)
FT      /tag= W
FT      /note= "AT5"
FT      complement (22653..23892)
FT      /tag= X
FT      /note= "KS5"
FT      complement (23992..46573)
FT      /tag= Y
FT      /note= "fkbB gene"
FT      complement (24163..24373)
FT      /tag= Z
FT      /note= "ACP4"
FT      complement (24997..26146)
FT      /tag= aa
FT      /note= "DH4 (inactive)"
FT      complement (26146..27430)
FT      /tag= ab
FT      /note= "AT4"
FT      complement (27430..28684)
FT      /tag= ac
FT      /note= "KS4"
FT      complement (28750..28960)
FT      /tag= ad
FT      /note= "ACP3"
FT      complement (29092..29740)
FT      /tag= ae
FT      /note= "KR3"
FT      complement (29869..31018)
FT      /tag= af
FT      /note= "DH3 (inactive)"
FT      complement (31018..32185)
FT      /tag= ag
FT      /note= "AT3"
FT      complement (32185..33439)
FT      /tag= ah
FT      /note= "KS3"
FT      complement (33505..33715)
FT      /tag= ai
FT      /note= "ACP2"
FT      complement (33823..34480)
FT      /tag= aj
FT      /note= "KR2"
FT      complement (34606..35749)
FT      /tag= ak
FT      complement (28750..28960)
FT      /tag= al
FT      /note= "DH2 (inactive)"
FT      complement (35749..37144)
FT      /tag= am
FT      /note= "AT2"
FT      complement (37145..38296)
FT      /tag= an
FT      /note= "KS2"
FT      complement (38371..38581)
FT      /tag= ao
FT      /note= "ACP1"
FT      complement (38677..39307)
FT      /tag= ap
FT      /note= "KR1"
FT      complement (39442..40609)
FT      /tag= aq
FT      /note= "DH1"
FT      complement (40609..41842)
FT      /tag= ar
FT      /note= "AT1"
FT      complement (41842..43093)
FT      /tag= as
FT      /note= "KS of extender module 1 (KS1)"
FT      complement (43144..43660)
misc_feature

```

```

FT      /tag= at
FT      /note= "ACP of loading domain"
FT      complement (43777..44629)
FT      /tag= au
FT      /note= "ER of loading domain"
FT      complement (44974..46573)
FT      /tag= av
FT      /note= "CoA ligase of loading domain"
FT      complement (46754..47788)
FT      /tag= aw
FT      /note= "fkbO gene"
FT      complement (47785..52272)
FT      /tag= ax
FT      /note= "fkbP gene"
FT      complement (52275..71465)
FT      /tag= ay
FT      /note= "fkbA gene"
FT      complement (52362..53576)
FT      /tag= az
FT      /note= "KS7"
FT      complement (53577..54716)
FT      /tag= ba
FT      /note= "AT7"
FT      complement (54717..55871)
FT      /tag= bb
FT      /note= "DH7"
FT      complement (56019..56819)
FT      /tag= bc
FT      /note= "ER7"
FT      complement (56943..57575)
FT      /tag= bd
FT      /note= "KR7"
FT      complement (57710..57920)
FT      /tag= be
FT      /note= "ACP7"
FT      complement (57990..59243)
FT      /tag= bf
FT      /note= "KS8"
FT      complement (59244..60398)
FT      /tag= bg
FT      /note= "AT8"
FT      complement (60399..61412)
FT      /tag= bh
FT      /note= "DH8 (inactive)"
misc_feature

```

Query Match 16.3% Score 44.6; DB 21; Length 77536;
 Best Local Similarity 53.1%; Pred. No. 0.13; Mismatches 84; Indels 0; Gaps 0;
 Matches 95; Conservative 0;

```

QY 96 tccatcgccacccgagaaagcgcgcattgacgcgcgcctctccacccctccctcc 155
DB 46351 TACCTGCGCGCCGCCGCCGCCGCGCTCCGCTCCGCTCAACCGGACGCCGCCGAC 46292
QY 156 gtcacccctcaccccaagagatttgaagaatacgcgcctacaagcgcgtcgatcgtg 215
DB 46291 GCCGAACCTCGCCCACTTCTTACGACTCGCGCGCACCTGCTGTCACGACGAACC 46232
QY 216 gagtcgcgcattgctcctcgcgcctagacccgcgcctcacccgaagatgcgtgacgcg 274
DB 46231 CACCTCGACACAGTGGCGCGGACCGGACCGACCGACCCACCTCTGCTCGTGGGCGCG 46173

```

```

RESULT 9
AAH44089
ID AAH44089 standard; cDNA; 1233 BP.
AC AAH44089;
XX
XX 12-SEP-2001 (first entry)
DE Oryza sativa peroxidase r2329 encoding cDNA SEQ ID NO:37.
XX
XX Oryza sativa; rice; peroxidase; POX; characteristic; gene expression;
KW

```


KM modification; plant; bacterial infection; Magnaporthe grisea; ss.
 OS Oryza sativa.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 34..1089
 FT /tag= a
 FT /product= "peroxidase r2329"
 XX
 XX WO200142475-A1.
 XX
 PD 14-JUN-2001.
 XX
 PF 08-DEC-2000; 2000WO-JP08728.
 XX
 PR 10-DEC-1999; 99JP-0352472.
 XX
 PA (NORO) JAPAN MIN AGRIC FORESTRY & FISHERIES.
 PI Ohashi Y, Mitsuhashi I, Sasaki T, Negamura Y, Ito H, Iwai T;
 PI Hiraga S;
 XX
 DR WPI: 2001-381695/40.
 DR P-PSDB; AAB99750.
 XX
 PT New set of rice peroxidase genes for analysis of peroxidase expression
 PT in rice under varying conditions and production of rice plants with
 PT desired characteristics
 XX
 PS Claim 1; Page 226-230; 258pp; Japanese.
 XX
 CC The present invention describes a set of peroxidase genes found in
 CC plants, especially rice, and their homologues, modified forms and
 CC fragments, where the sequences of the peroxidase genes in the set are
 CC given in AAH44071 to AAH44091. Also described are: (1) promoters for the
 CC control of the gene set; (2) the preparation of cassette vectors using the
 CC the genes and promoters; (3) analysis of plant characteristics using the
 CC peroxidase set by isolating RNA from the plant, binding the RNA to a
 CC membrane, mixing with a labelled peroxidase gene set, incubating, and
 CC detecting the label signal to show which genes in the set are expressed
 CC in the sample plant; and (4) DNA microarrays for peroxidase gene
 CC expression analysis. The set of genes are used for the analysis of the
 CC pattern of peroxidase gene expression in particular rice plants and
 CC their component tissues and under different environmental conditions,
 CC and modification of rice plants to provide desired specificities of
 CC peroxidase gene expression to impart particular characteristics to the
 CC plants such as response to bacterial infection by Magnaporthe grisea.
 XX
 SO Sequence 1233 BP; 228 A; 417 C; 360 G; 228 T; 0 other;
 XX
 Query Match 15.8%; Score 43.4; DB 22; Length 1233;
 Best Local Similarity 56.4%; Pred. No. 0.11; Mismatches 76; Indels 3; Gaps 1;
 Matches 102; Conservative 0;
 XX
 QY 92 ccaattatgcacagagaagccgcatgagccgctcctccaccctccccc 151
 DB 327 cggagtgatgcagacgcgcgacgagcggtgctgctgctgctgctgctg 386
 QY 152 ctccgcatctcaccacaaagattgaagaatgcgcgctcacaagcgctgagta 211
 DB 387 cgcgcgaatctcgtctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 446
 QY 212 cgtgagatgcgcatgctcctgcgctgacacgcgcgcgcgcgcgcgcgcgc 271
 DB 447 caaggtcccccgc 503
 QY 272 c 272
 DB 504 c 504
 RESULT 10

AAT94500
 ID AAT94500 standard; DNA: 701 BP.
 XX
 XX AAT94500;
 AC
 XX
 XX 20-MAR-1998 (first entry)
 XX
 DE GAL4 DNA binding domain fused to transcription activating VP16 DNA.
 XX
 XX GAL4; GAL4 DNA-binding domain; gene expression; plant;
 KM HSV VP16 transcriptional activation domain; plant enhancer sequence;
 KM glucuronidase; green fluorescent protein; ss.
 XX
 OS Chimeric - Saccharomyces cerevisiae.
 OS Chimeric - Human herpes simplex virus.
 XX
 FH Key Location/Qualifiers
 FT CDS 17..697
 FT
 FT misc_feature /tag= a
 FT misc_feature 1..458
 FT /tag= b
 FT /note= "GAL4 DNA-binding domain"
 FT misc_feature 459..701
 FT /tag= c
 FT /note= "HSV VP16 transcriptional activation domain"
 XX
 PN MO9730164-A1.
 XX
 PD 21-AUG-1997.
 XX
 PF 14-FEB-1997; 97WO-GH00406.
 XX
 PR 14-FEB-1996; 96GB-0003069.
 XX
 PA (MED1-) MEDICAL RES COUNCIL.
 PI Haseloff JP, Hodge S;
 PI
 DR WPI: 1997-425039/39.
 DR P-PSDB; AAW14300.
 XX
 PT DNA encoding GAL4 DNA binding domain optimised for plant expression
 PT - used for expressing genes in known pattern, for co-ordinating
 PT expression of several genes and for trapping plant enhancer
 PT sequences
 XX
 PS Example 1; Fig 1; 40pp; English.
 XX
 CC The present sequence represents a chimeric gene, encoding the 5'
 CC portion of the GAL4 DNA-binding domain and encoding in the 3' portion
 CC the transcriptional activation domain from HSV VP16. The A/T content of
 CC the GAL4 nucleotide sequence in this gene is less than 40% (contrasting
 CC with 59% in yeast). Reducing the A/T content optimises expression in
 CC plants. Constructs containing this nucleic acid can be used to express
 CC genes in a known pattern (temporal or spatial) in a plant, and to
 CC co-ordinate expression of several genes and, as traps, to isolate plant
 CC enhancer sequences. Particular applications are in studies of metabolic
 CC pathways and to modulate synthesis of e.g. dyes or lipids. Constructs
 CC containing sequences encoding the altered GAL4 DNA binding domain may
 CC include a reporter gene linked to a GAL4 response upstream activating
 CC sequence, particularly one encoding glucuronidase or green fluorescent
 CC protein.
 XX
 SO Sequence 701 BP; 147 A; 241 C; 194 G; 119 T; 0 other;
 XX
 Query Match 15.7%; Score 43; DB 18; Length 701;
 Best Local Similarity 59.3%; Pred. No. 0.13;
 Matches 73; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 QY 120 atgagc 179
 DB 263 atggaaccctccagacatcaaacgcctcgtcaccggtctcgtccagacacgtg 322

[illegible]

Key Location/Qualifiers
 CDS 10..1362
 /*tag= a
 /product= H_C peptide fragment

W0200067700-A2.

16-NOV-2000.

12-MAY-2000; 2000WO-US12890.

12-MAY-1999; 99US-0133865.

12-MAY-1999; 99US-0133866.

12-MAY-1999; 99US-0133867.

12-MAY-1999; 99US-0133868.

12-MAY-1999; 99US-0133869.

12-MAY-1999; 99US-0133870.

29-JUL-1999; 99US-0146192.

(USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.

Smith LA, Byrne MP, Middlebrook JL, Lagenotiere H;

WPI: 2001-016048/02.

P-PSDB: AAB04092.

New nucleic acids encoding the carboxy- or amino-terminal portions of

the heavy chain of botulinum neurotoxin of serotype A-G, useful as

vaccine against botulism

Claim 2; Fig 5a; 73pp; English.

Botulism neurotoxins are translated as a single 150 kDa polypeptide

chain and then posttranslationally nicked, forming a di-chain

consisting of a 100 kDa heavy chain and a 50 kDa light chain which

remain linked by a disulfide bond. Nucleic acids encoding the

carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy

chain of botulinum neurotoxin (BoNT) can be used in recombinant

expression vectors and expressed in transformed cells to produce

peptide antigens useful for eliciting an immune response to give

protective immunity against botulinum neurotoxin, which causes

botulism. The nucleic acids are expressible in a recombinant

organisms such as Escherichia coli or Pichia pastoris. The use

of recombinant nucleic acids are advantageous since it eliminates

the need to culture large quantities of hazardous toxin-producing

bacterium. Production yield from the genetically engineered product

is also high and cost of production is lower. The nucleic acids can

be derived from Clostridium botulinum serotypes A-G.

Sequence 1371 BP; 397 A; 416 C; 265 G; 293 T; 0 other;

Query Match 15.5%; Score 42.6; DB 22; Length 1371;

Best Local Similarity 50.2%; Pred. No. 0.18; Mismatches 104; Indels 0; Gaps 0;

Matches 105; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

61 ccttaacataacatgcatctccctacccattcatcgccacgcaagccgca 120

674 ccgaacataacatgcatctccctacccattcatcgccacgcaagccgca 733

121 tggagcgcggtctccacccctccctccctccctccctccctccctccctcc 180

734 agacataacatctcgttcacatctcgttcacatctcgttcacatctcgttc 793

181 agaaatcgccgctacagaagcgctgagtagtgcgagtagtgcgagtagtgc 240

794 gtaacgacctgagatcacagaagagtagtgcgagtagtgcgagtagtgc 240

241 gcaccggtctccacgcaagcatgcgctc 269

854 acatgtaacgcaactccagacagatcgctc 882

RESULT 15
 AA12317/C
 ID AA12317 standard; DNA; 771 BP.

AA12317;

08-OCT-1999 (first entry)

Neisseria meningitidis strain A complete ORF122 sequence.

Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.

Neisseria meningitidis.

W09924578-A2.

20-MAY-1999.

09-OCT-1998; 98WO-IB01665.

01-SEP-1998; 98GB-0019016.

06-NOV-1997; 97GB-0023516.

14-NOV-1997; 97GB-0024190.

18-NOV-1997; 97GB-0024386.

27-NOV-1997; 97GB-0025158.

10-DEC-1997; 97GB-0026147.

14-JAN-1998; 98GB-0000759.

(CHIR-) CHIRON SPA.

Grandi G, Masignani V, Pizza M, Rappelli R, Scarlato V;

WPI: 1999-327407/27.

P-PSDB: AAY38896.

Proteins from Neisseria meningitidis and N. gonorrhoeae useful for

diagnosis, treatment and prevention of infection

Claim 9; Page 440; 524pp; English.

Nucleotide sequences AA121972-212358 represent open reading frames

(ORFs) of Neisseria meningitidis and N. gonorrhoeae which encode

antigenic proteins (see AAY38499-Y38944). The antigenic proteins, their

fragments, their nucleic acids and antibodies are used for diagnosis,

prevention (as vaccines) or treatment of Neisseria infections,

such as meningitis, septicaemia and gonorrhea. Both organisms

are closely related. Fragments of the nucleic acids are useful

as hybridisation probes and antisense reagents.

Sequence 771 BP; 148 A; 207 C; 198 G; 213 T; 5 other;

Query Match 15.3%; Score 42; DB 20; Length 771;

Best Local Similarity 61.7%; Pred. No. 0.23;

Matches 66; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

165 accaagacgattggaagaatacgccgctacagaagcgctgagtagtgcgagtagtgc 224

191 ACACAAAGACGACACTCAAGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 132

225 atggtctcggtcgaagcagcgctcaccgcaagatgcgctga 271

131 GAATACATCGCGCATCGGTACCGGTTGACACATCACTTTTCATCGA 85

Search completed: July 3, 2002, 07:58:01
 Job time: 24436 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 07:58:01 ; Search time 796.36 Seconds
(without alignments)
830.042 Million cell updates/sec

Title: US-09-300-482-356
Perfect score: 385
Sequence: 1 caaccgacccaacgctgac.....acgctctggagttttttt 385

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*

- 1: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260.2	67.6	477	21	AAC35784 Zea mays DNA fragm
2	170.4	44.3	2629	17	AAT35903 Sequence encoding
3	170.4	44.3	2629	21	AA71193 Tobacco transketol
4	162.2	42.1	1681	21	AAC34305 Arabidopsis thalia
5	162.2	42.1	2614	21	AAC36886 Arabidopsis thalia
6	160.6	41.7	700	21	AAC34801 Arabidopsis thalia
7	155	40.3	1839	21	AAC46455 Arabidopsis thalia
8	155	40.3	2226	21	AAC43120 Arabidopsis thalia
9	116.2	30.2	1998	23	AAS54058 Pseudomonas aerugi

10	114.6	29.8	1998	22	AAS00429 Pseudomonas aerugi
11	114.6	29.8	1998	22	AAS00430 Pseudomonas aerugi
12	108.4	28.2	17775	23	AAS59506 Propionibacterium
13	93.8	24.4	2100	22	AAB66705 C glutamicum codin
14	93.8	24.4	2223	22	AAE71483 Corynebacterium gl
15	93.8	24.4	2350	22	AAS15846 C. glutamicum part
16	93.8	24.4	4108	22	AAE99988 DNA encoding Coryn
17	93.8	24.4	6995	22	AAE25332 Nucleotide sequenc
18	93.8	24.4	6995	22	AAE25332 C. glutamicum Opca
19	93.8	24.4	349980	22	AAH68529 C glutamicum codin
20	88.6	23.0	2944528	24	ABR03041 Listeria monocytog
21	86.4	22.4	2115	23	AAS56005 Salmonella typhi D
22	82.8	21.5	1992	22	AAH81452 Escherichia coli p
23	82.8	21.5	2143	22	AAD06950 ltkA gene encoding
24	82.8	21.5	6641	22	AAD06957 pME2 comprising ub
25	81.2	21.1	1698	23	AAS94417 DNA encoding novel
26	81.2	21.1	1989	23	AAS54405 staphylococcus aur
27	80.6	20.9	11443	19	AAV52182 Streptococcus pneu
28	80.4	20.9	738	23	AAS89780 DNA encoding novel
29	80.4	20.9	2004	23	AAS52511 E. coli DNA for ce
30	80.4	20.9	2637	23	AAS50091 DNA encoding novel
31	80.4	20.9	2637	23	AAS90091 DNA encoding novel
32	80.4	20.9	3051	23	AAS89789 DNA encoding novel
33	80.4	20.9	4275	23	AAS89787 DNA encoding novel
34	80.0	20.8	1953	18	AA774685 Staphylococcus aur
35	78.8	20.5	1971	23	AAS1643 Staphylococcus pne
36	78.2	20.3	1968	21	AA664393 Streptococcus pneu
37	78	20.3	1989	21	AA664393 DNA encoding a Sta
38	76.8	19.9	1998	23	AAS53437 Haemophilus Influe
39	74.4	19.3	48275	21	AAAB1501 N. meningitidis pa
40	74.4	19.3	349980	21	AAE21610 Neisseria meningit
41	74.4	19.3	349980	21	AAE21611 Neisseria meningit
42	74.4	19.3	837096	21	AAAB1489 N. meningitidis pa
43	71	18.4	5348	23	AAS90092 DNA encoding novel
44	67.4	17.5	1996	22	ABA89025 Escherichia coli p
45	67.4	17.5	16066	22	ABA89022 Escherichia coli p

ALIGNMENTS

RESULT 1	
AAC35784	
ID AAC35784 standard; DNA: 477 BP.	
XX	
AC AAC35784;	
XX	
DT 17-OCT-2000 (first entry)	
XX	
DE Zea mays DNA fragment SEQ ID NO: 11411.	
XX	
KW Hybridisation assay; genetic mapping; gene expression control;	
KW protein identification; signal transduction pathway; metabolic;	
KW pathway; promoter; termination sequence; corn; ss.	
XX	
OS Zea mays subsp. mays.	
XX	
PN EP1033405-A2.	
XX	
PD 06-SEP-2000.	
XX	
PF 25-FEB-2000; 2000EP-0301439.	
XX	
25-FEB-1999; 9905-0121825.	
PR 05-MAR-1999; 9905-0123180.	
PR 09-MAR-1999; 9905-0123548.	
PR 23-MAR-1999; 9905-0125788.	
PR 25-MAR-1999; 9905-0126264.	
PR 29-MAR-1999; 9905-0126785.	
PR 01-APR-1999; 9905-0127462.	
PR 06-APR-1999; 9905-0128234.	
PR 08-APR-1999; 9905-0128714.	
PR 16-APR-1999; 9905-0129845.	

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0138769.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141284.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

5

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160960.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161362.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 67.6%; Score 260.2; DB 21; Length 477;
Best Local Similarity 95.4%; Pred. No. 3.6e-56;
Matches 268; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 1 caaccgagaccgaagcctacatcattggttgaggaccggtcccgagctcgagatccgcg 60
Db 197 cgaccggaacaagccctgacccatctgagtgaccggtcccgagctcgagatccgcg 256
OY 61 gcaatgcgagccgagcagctgaggaagggaggaagcagctccgcgctcgtctcgtct 120
Db 257 ccaagggccgagcagcagctgaggaagggaggaagcagctccgcgctcgtctcgtct 316
OY 121 cctgggaactctttgatgagcagctcgatgatacaagagagagcgtccctccctcgagc 180
Db 317 cctgggaactctttgatgagcagctcgatgatacaagagagagcgtccctccctcgagc 376
OY 181 tcacgagcagatcagcagctcagagccgagctccacatccgctcgagagatcagcag 240
Db 377 tcacgagcagatcagcagctcagagccgagctccacatccgctcgagagatcagcag 436
OY 241 cccaaggaagggcattgagcagcagcaagttcgagcgag 281
Db 437 cccaaggaagggcattgagcagcagcaagttcgagcgag 477

RESULT 2
AAT35903
ID AAT35903 standard; cDNA: 2629 BP.

XX AAT35903;

XX AC 29-JAN-1997 (first entry)

XX DT Sequence encoding tobacco transketolase.

XX DE Transketolase; herbicide; Nicotiana tabacum; identification;
XX KW screening; ss.

XX OS Nicotiana tabacum.

XX XX Key Location/Qualifiers
XX FT 60..2291
XX CDS /tag= a
XX /product= Transketolase.

XX EP723017-A2.
XX PN 24-JUL-1996.
XX PD

XX 13-JAN-1996; 96EP-0100458.
XX 23-JAN-1995; 95DE-4001906.
XX (SCHM/) SCHMIDT R.
XX (BADI) BASF AG.
XX Schmidt R, Sonnewald U, Stitt M;
XX WPI: 1996-335478/34.
XX P-PSDB; AAM03319.
XX New transketolase from tobacco and related DNA - useful for
XX screening for inhibitors which are potential herbicides
XX Claim 5; Page 9-12; 26pp; German.
XX The transketolase from tobacco is useful for identifying herbicides
XX for their ability to inhibit the transketolase. The transketolase
XX can be produced in pure form by cloning the gene.
XX Sequence 2629 BP; 686 A; 595 C; 620 G; 728 T; 0 other;

Query Match 44.3%; Score 170.4; DB 17; Length 2629;
Best Local Similarity 67.4%; Pred. No. 1.8e-33;
Matches 240; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

OY 4 ccggaccacgaagcctacatcattggttgaggaccggtcccgagctcgagatccgcgag 63
Db 1936 ctgggaacaacaacctgattgattgattgattgattgattgattgattgattgattgatt 1995
OY 64 atcgagccgagcagctgaggaagggaggaagcagctccgcgctcgtctcgtctcct 123
Db 1996 aggcgctgattgattgattgattgattgattgattgattgattgattgattgattgatt 2055
OY 124 ggggaactctttgatgagcagctcgatgatacaagagagagcgtccctccctcgagc 183
Db 2056 ggggaactctttgatgagcagctcgatgatacaagagagagcgtccctccctcgagc 2115
OY 184 cagcagagcagcagctcagagccgagctccacatccgctcgagagatcagcagc 243
Db 2116 cagcagagcagcagctcagagccgagctccacatccgctcgagagatcagcagc 2175
OY 244 aaggaagggcattgagcagcagcaagttcgagcgagtgctcctcgaggagcagc 303
Db 2176 aaggaagggcattgagcagcagcaagttcgagcgagtgctcctcgaggagcagc 2235
OY 304 aggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 359
Db 2236 aggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2291

RESULT 3
AAAT1793
ID AAAT1793 standard; cDNA: 2629 BP.

XX AAAT1793;

XX AC 08-JAN-2001 (first entry)

XX DT Tobacco transketolase cDNA.

XX DE Transketolase; tobacco; transgenic plant; tocopherol; Vitamin K; lignin;
XX KW chlorophyll; lignin; aromatic amino acid; soya; canola; barley; oat;
XX KW wheat; rape; maize; sunflower; ss.

XX OS Nicotiana tabacum.

XX XX Key Location/Qualifiers
XX FT 60..2291
XX CDS /tag= a

XX EP723017-A2.
XX PN 24-JUL-1996.
XX PD

PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142970.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147182.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156536.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 42.1%; Score 162.2; DB 21; Length 1681;
Best Local Similarity 66.4%; Pred. No. 1.9e-31;
Matches 233; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 4 ccggaccgaacgcctgactatcttggtgtggcaccgcgctccgagcttgagatcgcgagca 63
DB 1075 caggcaacaacccgagttgacttcttgattggaactggtcttcgactagagatcttgac 1134
QY 64 atcgcgccgacgagcttgaggagggggaagacggtcgcgctcgttcgtctcct 123
DB 1135 agctcgcgaggtgctcaggaagaagacggaacccgttagagttgttcttcgtgtgct 1194
QY 124 gggactctttgatgagcagtcgtagtgatcaagggagagcgctccctccgcagcgta 183
DB 1195 gggactatttgacgagcaatcagatgatacaaggagagtggtgtgcatcagatgat 1254
QY 184 cagcgagatcagatcagagccggtgctcactcttcgttcgagcagaagactcgagccc 243
DB 1255 cagctagagttagacttgaacagcacttcgacttctcggtgggaaagattgttgagga 1314
QY 244 aagcgaagccttgagcctgcagacaagttcggcgagtgctcccgcgagcgtctaca 303
DB 1315 aaggaagtcattcgatattcaatcattcagcgagcagcagcagcacttaactaca 1374

OY 304 aggaagacgcacacgcgtgagagacatcatgccaactgccaagacgtt 354
||||| | | ||||| | | | | |
Db 1375 aggaattcgtatcacgcgttgaaagctgtgtgatgcgcgaagcatctt 1425

RESULT 5

ID AAC36886 standard; DNA; 2614 BP.

AAC36886;

17-OCT-2000 (first entry)

Arabidopsis thaliana DNA fragment SEQ ID NO: 15417.

XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EPI033405-A2.

PD 06-SEP-2000.

PE 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139762.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.


```
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157513.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 42.1%; Score 162.2; DB 21; Length 2614;
Best Local Similarity 66.4%; Pred. No. 2,1e-31;
Matches 233; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
```

```
OY 124 gggactctttgatgagcagtcgcatgagtaacaggagacgctctccctgcgcagctca 183
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2128 gggactatttgacgagaaatcagatgatacaaggagagtgcttgcacatcgatgat 2187
OY 184 cagcgagagatcagatcagcagcggtccactctcgtctgcagagaagctcgagccc 243
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2188 cagctagatttagcattgaacagcttcgactctcgtatgggaaagatttgttgaaggca 2247
OY 244 aagcgagcgcatctggatcgcagcaagttcgcgcgagtcgctccctgcggagcatctaca 303
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2248 aaggaagatccattggtattatcattcgcagcagcagcaccacacttactctaca 2307
OY 304 aggagtcgcatcacccgctggagagacatcattgcaactgcgaagacttt 354
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2308 aggagtttgatcacccgttgagacgtgtgttgatgcggccaatcattct 2358

RESULT 6
AAC34801
ID AAC34801 standard; DNA; 700 BP.
XX
AC AAC34801;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 7936.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
```


PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145291.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149368.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151458.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0156596.
PR 05-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 12-OCT-1999; 99US-0158232.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 23-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 28-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161933.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match	40.3%	Score 155;	DB 21;	Length 1839;
Best Local Similarity	65.4%;	Pred. No. 1.2e-29;		
Matches 227; Conservative	0;	Mismatches 120;	Indels 0;	Circ 0

[illegible]

RESULT	8
AAC43120	
ID AAC43120	standard; DNA; 2226 BP.

AAC43120;

17-OCT-2000 (first entry)

Arabidopsis thaliana DNA fragment SEQ ID NO: 38100.

Hybridisation assay; genetic mapping; gene expression control; protein identification; signal transduction

metabolic pathway; promoter; termination sequence; ss

Arabidopsis thaliana

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

43-HEB-1999; 99US-0121825.
05-MAB-1999; 99US-0122190

03 MAR-1999;	99US-0123548.
23-MAR-1999;	99US-0125788

29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132485.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141847.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158363.
PR 13-OCT-1999; 99US-0159293.

QY 311 cggcatcacggtgagagcatcatgtcaactgccaaag 347
|||||
Db 1944 cggcttcacctggaacaagcttcgcyggtgcgcgaag 1980

RESULT 10
AAS00429
ID AAS00429 standard; DNA; 1998 BP.

AC AAS00429;

DT 11-MAY-2001 (first entry)

DE Pseudomonas aeruginosa transketolase tkta gene.

XX Transketolase; tkta; microbial growth; antimicrobial; antibacterial;
KW bacterial infection; gastric ulcer; chronic lung infection; gastritis;
KM cystic fibrosis; microbial disease; Helicobacter pylori infection;
KW gastrointestinal carcinoma; endocarditis; bacteraemia; pneumonia;
KM osteomyelitis; otitis media; urinary tract infection; folliculitis; ds.

OS Pseudomonas aeruginosa.

FH Key Location/Qualifiers

FT CDS

FT 1.1998
/*tag= a
/product= "tkta protein"

PN W0200114523-A1.

PD 01-MAR-2001.

PE 16-AUG-2000; 2000WO-US22324.

PR 24-AUG-1999; 99US-0382106.

PA (SMIK) SMITHKLINE BEECHAM CORP.

PA (SMIK) SMITHKLINE BEECHAM PLC.

PI Huang J, Jiang X, McDevitt D, Van Horn S;

DR WPI: 2001-218437/22.

DR P-PSDB; AAU00425.

XX New Pseudomonas aeruginosa tkta polypeptide and polynucleotide, useful
PT for screening antibacterial compounds for treating microbial diseases,
PT and as diagnostic reagents for diagnosing or prognosing bacterial
PT infection -

XX Claim 2; Page 3; 40pp; English.

XX The present sequence encodes for a novel Pseudomonas aeruginosa
CC transketolase tkta protein which is related to other proteins of the
CC transketolase family. The tkta protein is essential for the growth
CC and/or survival of P. aeruginosa or any organism that may possess it.
CC This is shown in particular by a temperature sensitive (ts) mutant
CC ts-92 (AAU00426) which has an amino acid substitution of Arg to His at
CC residue 204. The tkta polypeptides are useful as antimicrobial
CC particularly antibacterial compounds and identifying membrane bound or
CC soluble receptors. The tkta polynucleotides can be used as diagnostic
CC reagents. Both the tkta polypeptides and polynucleotides are useful for
CC diagnosing or prognosing a disease or susceptibility to a disease in a
CC human (e.g. bacterial infection) and for assessing the binding of small
CC molecule substrates and ligands in cells and chemical libraries. Diseases
CC caused by P. aeruginosa are difficult to treat because of antibiotic
CC resistance and untreatable strains often found in patients with chronic
CC lung infections e.g. cystic fibrosis. The agonists and antagonists to
CC the tkta polypeptides and/or polynucleotides are useful for treating
CC microbial diseases such as endocarditis, bacteraemia, pneumonia, otitis,
CC osteomyelitis, folliculitis and urinary tract and wound infections.
CC Helicobacter pylori infections, reducing the risk of H. pylori-induced
CC cancers such as gastrointestinal carcinoma and also preventing/treating
CC gastric ulcers and gastritis.

XX SQ Sequence 1998 BP; 365 A; 710 C; 636 G; 287 T; 0 other:

Query Match 29.8%; Score 114.6; DB 22; Length 1998;
Best Local Similarity 58.8%; Pred. No. 1.8e-19;
Matches 198; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 11 caagctgacacatcttggttggcaccggtcccgagctgagatcgcgggcaatgcggc 70
|||||

Db 1644 cgagccggaactgatactcgtacgcacccggttcggaagtcggcgtgcgcgccta 1703
|||||

QY 71 cgaagcgtcgaaggagggggaagacggtccggtcgttcgttcgttcctcctggaaact 130
|||||

Db 1704 cgacaagctcagcgagccagcgccgaaggttcggtggtatcgtatgcattaccacgcgt 1763
|||||

QY 131 ctctgatgagcagtcggatgagtaacaagagagcgtccctccctcccgacgtcacagcgag 190
|||||

Db 1764 ctacgagcagcagaagcaggtcctacaagcagtcgtcgtccggttggaagtcgcgcgcg 1823
|||||

QY 191 gatcagcatcgagcgccgggtccactctcgtctggcagaagtaagtcggagcccaagcga 250
|||||

Db 1824 catcgccatcgagcgccgccatgcgcgactactgttacaagtaagtcgtcgcagcggcg 1883
|||||

QY 251 ggcattggcatcgacaagtcgcgcgagtgctccctcgcggagacatacaaggagta 310
|||||

Db 1884 catcatcgcatgacacagcttcgcgcgagtcgcgcgcgcgcgcgcgtgttcgagcact 1943
|||||

QY 311 cggcatcacggtgagagcatcatgtcaactgccaaag 347
|||||
Db 1944 cggcttcacctggaacaagcttcgcyggtgcgcgaag 1980

RESULT 11

AAS00430
ID AAS00430 standard; DNA; 1998 BP.

AC AAS00430;

DT 11-MAY-2001 (first entry)

DE Pseudomonas aeruginosa transketolase tkta mutant ts-92 DNA.

XX Transketolase; tkta; microbial growth; antimicrobial; antibacterial;
KW bacterial infection; gastric ulcer; chronic lung infection; gastritis;
KW cystic fibrosis; microbial disease; Helicobacter pylori infection;
KW gastrointestinal carcinoma; mutant; mutain; endocarditis; bacteraemia;
KW pneumonia; osteomyelitis; otitis media; urinary tract infection;
KW folliculitis; ds.

OS Pseudomonas aeruginosa.

OS Synthetic.

FH Key Location/Qualifiers

FT CDS

FT 1.1998
/*tag= a
/product= "tkta mutant ts-92"

FT mutation

FT replac (611,6)
/*tag= b

PN W0200114523-A1.

PD 01-MAR-2001.

PE 16-AUG-2000; 2000WO-US22324.

PR 24-AUG-1999; 99US-0382106.

PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PI Huang J, Jiang X, McDevitt D, Van Horn S;
XX


```
QY 75 gagctgaggaagaggggaagacggtccggtctcgtctcgtctccttgggaactctt 134
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2008 gctctggaagctgagggcggttgacgctcggttgcttccgttcccttgcatgtgttc 2067
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 135 gatgagcagtcgagtgataaagaagagcgctccctcctgacggtcacagcgagatc 194
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2068 caggaagcagagcagcagatcatcagtcggtctcgtcgtcagcgttgaccgctcgttg 2127
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 195 agcatcgagcccggtccactctcgcttgagagaagtaagtcgagagcccaaggcaagcc 254
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2128 tctgttgaagctggcatcgcatctggttaccgcttcttggcaaccagggccgtgct 2187
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 255 attggcatcgacaagttcgccgagtgctctcgtcgggacgacatcaagaagtaagcgc 314
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2188 gtcccttgagcaccttcggtctcgtcgattaccagaccctgttgagaagttcggc 2247
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 315 atcacccgtgagagcatcatctgcaactgccaagact 351
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2248 atcacaccgatgcagtcgtggcagcgccaagact 2284
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: JULY 3, 2002, 07:58:10
Job time: 24445 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 08:11:28 ; Search time 185.98 Seconds
(without alignments)
508.490 Million cell updates/sec

Title: US-09-300-482-356

Perfect score: 385
Sequence: 1 caaccgcagcaagcctgac.....acggtctggagttttttt 385

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCPMUS.COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	170.4	44.3	2629	2 US-09-012-030-1	Sequence 1, Appl1
2	170.4	44.3	2629	2 US-08-590-454-1	Sequence 1, Appl1
3	114.6	29.8	1998	4 US-09-382-106-1	Sequence 1, Appl1
4	82.8	21.5	2143	4 US-09-064-693A-18	Sequence 18, Appl1
5	82.8	21.5	6641	4 US-09-064-693A-25	Sequence 25, Appl1
6	78.2	20.3	1968	4 US-09-298-724-1	Sequence 1, Appl1
7	76.8	19.9	1995	5 PCT-US96-05320A-1025	Sequence 1025, Ap
8	63	16.4	1738	4 US-08-858-307A-35	Sequence 35, Appl1
9	51.6	13.4	1872	2 US-08-743-637B-17	Sequence 17, Appl1
10	51.6	13.4	1872	3 US-08-526-840B-17	Sequence 17, Appl1
11	44	11.4	1284	4 US-09-327-487A-4	Sequence 4, Appl1
12	43	11.2	1281	4 US-09-327-487A-3	Sequence 3, Appl1
13	43	11.2	2872	4 US-09-327-487A-2	Sequence 2, Appl1
14	41.2	10.7	4195	1 US-08-360-011-1	Sequence 1, Appl1
15	41.2	10.7	4195	3 US-08-901-710-1	Sequence 1, Appl1
16	41.2	10.7	4416	3 US-08-795-430-1	Sequence 1, Appl1
17	41.2	10.7	4425	1 US-08-222-616-31	Sequence 31, Appl1
18	41.2	10.7	4425	5 PCT-US95-04228-31	Sequence 31, Appl1
19	41.2	10.7	4795	1 US-08-340-011-3	Sequence 3, Appl1
20	41.2	10.7	4795	3 US-08-901-710-3	Sequence 3, Appl1
21	41.2	10.7	9108	5 PCT-US95-04228-45	Sequence 45, Appl1
22	41	10.6	1693	3 US-09-320-878-23	Sequence 23, Appl1
23	37	9.6	71989	4 US-09-443-501A-2	Sequence 2, Appl1
24	36.6	9.5	4403765	4 US-09-103-840A-2	Sequence 1, Appl1
25	36.6	9.5	4411529	4 US-09-103-840A-1	Sequence 1, Appl1
26	36.2	9.4	1931	2 US-09-130-114-2	Sequence 2, Appl1
27	35.6	9.2	1136	1 US-08-106-981-5	Sequence 5, Appl1

28	35.6	9.2	1431	1 US-08-254-357-1	Sequence 1, Appl1
29	35.6	9.2	1632	1 US-07-959-941-1	Sequence 1, Appl1
30	35.6	9.2	1632	1 US-08-259-924-1	Sequence 1, Appl1
31	35.6	9.2	4983	1 US-08-472-358-1	Sequence 1, Appl1
32	35.6	9.2	4983	5 PCT-US92-05786A-1	Sequence 1, Appl1
33	35.4	9.2	68750	3 US-09-335-40786A-1	Sequence 1, Appl1
34	35.4	9.2	68750	4 US-09-568-102-1	Sequence 1, Appl1
35	35.4	9.2	68750	4 US-09-567-969-1	Sequence 1, Appl1
36	35.4	9.2	68750	4 US-09-568-480-1	Sequence 1, Appl1
37	35.4	9.2	68750	4 US-09-568-486-1	Sequence 1, Appl1
38	35.4	9.2	68750	4 US-09-568-472-1	Sequence 1, Appl1
39	35.2	9.1	711	4 US-09-651-941-22	Sequence 18, Appl1
40	35.2	9.1	711	4 US-09-651-941-22	Sequence 22, Appl1
41	35.2	9.1	12508	4 US-09-651-941-1	Sequence 1, Appl1
42	35.2	9.1	12523	4 US-09-651-941-1	Sequence 1, Appl1
43	34.8	9.0	1026	4 US-09-275-384B-4	Sequence 4, Appl1
44	34.8	9.0	1029	4 US-09-449-437A-1	Sequence 1, Appl1
45	34.8	9.0	2238	1 US-08-742-011-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-012-030-1
Sequence 1, Application US/09012030
Patent No. 5912169
GENERAL INFORMATION:
APPLICANT: SCHMIDT, Ralf-Michael, STITT, Marc, SONNEWALD,
APPLICANT: Uwe
TITLE OF INVENTION: Transketolase
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Keil & Weinlauf
STREET: 1101 Connecticut Avenue
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM AT-compatible, 80486 processor
OPERATING SYSTEM: MS-DOS version 6.0
SOFTWARE: Wordperfect version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012.030
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/590,454
FILING DATE: 22-JAN-1996
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2629 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE: Nicotiana
ORGANISM: Nicotiana
FEATURE:
NAME/KEY: CDA
LOCATION: 60..2289
US-09-012-030-1

Query Match 44.3%; Score 170.4; DB 2; Length 2629;
Best Local Similarity 67.4%; Pred. No. 6.8e-37;
Matches 240; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
OY 4 ccgcgcacgaagcctgacatctgtgtggcagcgtccgcgagctgagatccgggca 63
| ||||| ||| ||||| ||||| | || || ||||| ||| || || || || ||


```

Db 1936 CTGGCAACAACCTGATGTCATTTTGATTGGTACTGGCTCAGAGTAGAATTCGTCTCA 1995
QY 64 atcgccgacgagcttgaggaagaagggaagcgcgcgtcgtcgttcgttcgtcct 123
Db 1996 AGGCGTGATGACCTCAGGAAAGAGAAAGACAGTAGAGTGTGTTCCCTTGTTGTT 2055
QY 124 gggaaactcttgatgagcaatcgatagtaagaagagagcgtccctcccgagagtcga 183
Db 2056 GGGAGCTTTTGAAGAACATCAGCCGACTACAGGAAGTGCTTCATCATCTGTTA 2115
QY 184 cagcgagatcatgacatcgagccggtccacctcgtcgtcgagaaagtaacgtcgagacc 243
Db 2116 CAGCTAGAGTTAGCATTTAGAGCCGATCCACATTTGGGTGGAGAAATATGTCGATCAA 2175
QY 244 aagcgaaagccatgcatcgacagcaagttcgcgagagtcgtccctcgccgagacatcaca 303
Db 2176 AGGGGAAGCCCATCGGAATTGACAGATGGGTGCGCATGCCCCCTCGGAAATATATACA 2235
QY 304 aggagtaagcagcatcagcgttgagagagcatcattgcaactgcaagagctttaagag 359
Db 2236 AGGAGTAGGAATTACAGCAGAGCGCTGTGTGCTGACGCTAAACAAGTTTCTTAG 2291

```

```

RESULT 2
US-08-590-454-1
; Sequence 1, Application US/08590454
; Patent No. 592535
; GENERAL INFORMATION:
; APPLICANT: SCHMIDT, Ralf-Michael, STITT, Marc, SONNEWALD,
; APPLICANT: Dwe
; TITLE OF INVENTION: Transketolase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kell & Weinkauff
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM AT-compatible, 80486 processor
; OPERATING SYSTEM: MS-DOS version 6.0
; SOFTWARE: Wordperfect version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/590,454
; FILING DATE: 22-JAN-1996
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2629 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Nicotiana
; FEATURE:
; NAME/KEY: CDA
; LOCATION: 60..2289
; US-08-590-454-1

```

```

Query Match 44.3%; Score 170.4; DB 2; Length 2629;
Best Local Similarity 67.4%; Pred. No. 6.8e-37;
Matches 240; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
QY 4 ccggacaaagcctgacatcatttggttgccacggctccgagctcgagatcgcggagca 63
Db 1936 CTGGCAACAACCTGATGTCATTTTGATTGGTACTGGCTCAGAGTAGAATTCGTCTCA 1995
QY 64 atcgccgacgagcttgaggaagaagggaagcgcgcgtcgtcgttcgttcgtcct 123

```

```

Db 1996 AGGCTGCTGATGAACTCAGGAAGAAAGACGATGAGATTGTTCCCTTGTTGTT 2055
QY 124 gggaaactcttgatgagcaatcgatagtaagaagagcgtccctcccgagagtcga 183
Db 2056 GGGAGCTTTTGAAGAACATCAGCCGACTACAGGAAGTGCTTCCTTCATCATCTGTTA 2115
QY 184 cagcgagatcatgacatcgagccggtccacctcgtcgtcgagaaagtaacgtcgagacc 243
Db 2116 CAGCTAGAGTTAGCATTTAGAGCCGATCCACATTTGGGTGGAGAAATATGTCGATCAA 2175
QY 244 aagcgaaagccatgcatcgacaaagttcgcgagagtcgtccctcgccgagacatcaca 303
Db 2176 AGGGGAAGCCCATCGGAATTGACAGATGGGTGCGCATGCCCCCTCGGAAATATATACA 2235
QY 304 aggagtaagcagcatcagcgttgagagagcatcattgcaactgcaagagctttaagag 359
Db 2236 AGGAGTAGGAATTACAGCAGAGCGCTGTGTGCTGACGCTAAACAAGTTTCTTAG 2291

```

```

RESULT 3
US-09-382-106-1
; Sequence 1, Application US/09382106
; Patent No. 6221631
; GENERAL INFORMATION:
; APPLICANT: Huang, Jianzhong
; APPLICANT: Huang, Xime
; APPLICANT: McDevitt, Damien
; APPLICANT: Van Horn, Stephanie
; TITLE OF INVENTION: EKA
; FILE REFERENCE: GM10236
; CURRENT APPLICATION NUMBER: US/09/382,106
; CURRENT FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-382-106-1

```

```

Query Match 29.8%; Score 114.6; DB 4; Length 1998;
Best Local Similarity 58.8%; Pred. No. 6e-22;
Matches 198; Conservative 0; Mismatches 139; Indels 0; Gaps 0;
QY 11 caagcctgacatcattggttggtgacccggtctcgagctcgagagatcgggcaatgcggc 70
Db 1644 cgagccggaactgatactcgtatcgccacccggttcggaagtcggcctgacggccta 1703
QY 71 cgacgagcttgaggaaggaggaagacggtccgctgctcgtctcgttcgttccttggaact 130
Db 1704 cgacaagctgagcagcagggccgcaaggtcggctggtatcgtatgatacagcaagcgt 1763
QY 131 ctctgagtagcagctgagatgagtaagaagagcgttcctccctcccgagctacaagcgag 190
Db 1764 ctacgagcagcagcagcagctccacaagaagcagtcgtcgtccggtgagagtcggtcgcg 1823
QY 191 gatcagcatcagagccgggttcacatctcgtcgtcgagaaagtaagtcagcagcgaagaa 250
Db 1824 catcgccatcgagggccgccatcgccgactacgtgtaagaatcgtcgttcctcgagggcg 1883
QY 251 ggcacatgcatcgacaagtctcgcgcgagtgctcctcccgagagcatcacaagagta 310
Db 1884 catcatcgagcagacaaagcttcggagagtcggtcggtcccgcccggtcttcgactt 1943
QY 311 cggcatcagcgttgagagagcatcattcgacaactgccaag 347
Db 1944 cggctcaaccctggacaagctcctcggtcggtggcgag 1980

```

```

RESULT 4
US-09-064-693A-18

```


Sequence 18, Application US/09064693A
Patent No. 6210937
GENERAL INFORMATION:
APPLICANT: Ward, Thomas E.
TITLE OF INVENTION: DEVELOPMENT OF GENETICALLY
ENGINEERED BACTERIA FOR PRODUCTION
OF A SPECIFIC PLASTICS PRECURSOR
TITLE OF INVENTION: OF A SPECIFIC PLASTICS PRECURSOR
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: W. Gary Goodson
ADDRESS: INBEL--Lockheed Martin Idaho
ADDRESSEE: Technologies Co.
STREET: P.O. Box 1625
CITY: Idaho Falls
STATE: Idaho
COUNTRY: USA
ZIP: 83415-3810
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
COMPUTER: Toshiba Satellite Pro T2150CDS
OPERATING SYSTEM: Windows95
SOFTWARE: Word perfect 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/064,693A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: W. Gary Goodson
REGISTRATION NUMBER: 22,387
REFERENCE/DOCKET NUMBER: LIT-PI-296
TELECOMMUNICATION INFORMATION:
TELEPHONE: (208)526-9469
TELEFAX: (208)526-8339
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2143 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-064-693A-18

	Query Match	21.5%:	Score 82.8:	DB 4:	Length 2145:
	Best Local Similarity	53.0%:	Pred. No.2.le-13:		
	Matches 177:	Conservative	0:	Mismatches 157:	Indels 0:
				Gaps	0:
QY	13	agccctgacatcatcttggtcttggtgcacggcgtcccgactcgagactgagatctcgcggaatctgcgcccg	72		
Db	1726	AGCGGAACCTATTTTCATCCCTACACGGTTCCAGAGTTGAACCTGGCTGTGCTGCTTACG	1785		
QY	73	acgagctgaggaagaagagggaagacgctcgctgcgtctgcttgccttccttgygaactct	132		
Db	1786	AAAACTCATTCTCCGGAAGCGCTGAAGACGCCGCTGGTTCATATGCTACCGACGAT	1845		
QY	133	tttgttgagcaatctgcatgaatgacaaagsgagcgtctctccctcgcaagctccacaggaaga	192		
Db	1846	TTGACAAACAGATGCTCTTACCTGATCCGTACTCTCCGAACGGTCTTACTGTGACGCG	1905		
QY	193	tcaagcatcgaaagccgggtctccactctcgctctggacaaagtacgtctgagagcccaaggaag	252		
Db	1906	TTGCTGTAGAGCGGGTATTCTGACTACTGTGTACAAATATATTGTGGCCTGAACGGTGGTA	1965		
QY	253	ccatttgcatctgacaagtcttcggtcgagtgctctctcgtccgggaagcatctacaagaagtagc	312		
Db	1966	TCTGCGGATATGACCACTTCGGTGAATCTGCTCCGGCAGAGCTGCTGTTGAAGAGTTTCG	2025		
QY	313	gcataccgcctgtagagatcatcttgcacttgcgca	346		
Db	2026	GCTTCACATGTTGATTAACGTTGTGGGAAGCAAA	2059		

```

1      RESULT      5
2      US-09-064-693A-25
3      ; Sequence 25, Application US/09064693A
4      ; Patent No. 6210937
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Ward, Thomas E.
7      ; TITLE OF INVENTION: DEVELOPMENT OF GENETICALLY
8      ; TITLE OF INVENTION: ENGINEERED BACTERIA FOR PRODUCTION
9      ; TITLE OF INVENTION: OF A SPECIFIC PLASTICS PRECURSOR
10     ; NUMBER OF SEQUENCES: 26
11     ; CORRESPONDENCE ADDRESS:
12     ; ADDRESSEE: W. Gary Goodson
13     ; ADDRESSEE: INEEL--Lockheed Martin Idaho
14     ; ADDRESSEE: Technologies Co.
15     ; STREET: P.O. Box 1625
16     ; CITY: Idaho Falls
17     ; STATE: Idaho
18     ; COUNTRY: USA
19     ; ZIP: 83415-3810
20     ; COMPUTER READABLE FORM:
21     ; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
22     ; COMPUTER: Toshiba Satellite Pro T2150CDS
23     ; OPERATING SYSTEM: Windows95
24     ; SOFTWARE: Word Perfect 7.0
25     ; CURRENT APPLICATION DATA:
26     ; APPLICATION NUMBER: US/09/064,693A
27     ; FILING DATE:
28     ; CLASSIFICATION: 435
29     ; PRIOR APPLICATION DATA:
30     ; APPLICATION NUMBER:
31     ; FILING DATE:
32     ; ATTORNEY/AGENT INFORMATION:
33     ; NAME: W. Gary Goodson
34     ; REGISTRATION NUMBER: 22,387
35     ; REFERENCE/DOCKET NUMBER: LIT-PI-296
36     ; TELECOMMUNICATION INFORMATION:
37     ; TELEPHONE: (208)526-9469
38     ; TELEFAX: (208)526-8339
39     ; INFORMATION FOR SEQ ID NO: 25:
40     ; SEQUENCE CHARACTERISTICS:
41     ; LENGTH: 6641 base pairs
42     ; TYPE: nucleic acid
43     ; STRANDEDNESS: double
44     ; TOPOLOGY: linear
45     ;
46     ; US-09-064-693A-25

```

Query Match	21.5%;	Score 82.8;	DB 4;	Length 6641;
Best Local Similarity	53.0%;	Pred. No. 3e-13;		
Matches 177;	Conservative 0;	Mismatches 157;	Indels 0;	Gaps 0;
OY	13	agccctgacatcaltggtgtctggcaccgcgcgtccgcgactgagatcgcggygcaatgcgcgcg	72	
Db	3768	AGCGGAACTATATTTTTCACTCCGCTACCGCGTTGAAAGTTGAACCTGGCTTCTGCTCAAG	3827	
OY	73	acgagctgagaaagagagggaagaacgcgtccgcgttcgtctcgtctccttggaaactct	132	
Db	3828	AAAACTACTGTCGGCAAGCGCGTGAAGCGCGCGCGTGTCATGTCGTACCAACGCAT	3887	
OY	133	ttgatgagcaatcgtgatagtatcaagaaggagcgtccctccgcgcgaogtcaaaacgggga	192	
Db	3888	TTGAAACACAGAGATGCTCCTTACCGTGATCCGTAACTCCCAAGCGGTTACTCCACGCG	3947	
OY	193	tcaagatcggagccggggtccactctcggctcgtgcggaagtacgtcggagcccaaggcaag	252	
Db	3948	TTGCTGTGAAGCGGGATTTGCTGACTACTGTGACAAAGATGTTGGCTGAACGGTCTA	4007	
OY	253	ccattggatcgcacaagttcgcgcgagttctctcgcgcggagacatctacaagaagtagc	312	
Db	4008	TGCTGCGATATACCACTTCGCGTGAATGCTGCCGCAAGACTCTCTTTGAAAGATTGC	4067	
OY	313	gcataccgctggagagcatcaltgcgaactgccaa	346	

Db 4068 GCTTCACGTGTGATTAACCTTGTGGCAAGCAAA 4101

RESULT 6

US-09-298-724-1
Sequence 1, Application US/09298724
Patent No. 6326167
GENERAL INFORMATION:
APPLICANT: Yu, Jun
APPLICANT: McDevitt, Damien
APPLICANT: Lonetto, Michael A.
APPLICANT: Marra, Andrea
APPLICANT: Holden, David A.
APPLICANT: Ingraham, Karen A.
APPLICANT: Ray, Jennifer
APPLICANT: Chalker, Allison F.
APPLICANT: Holmes, David J.
APPLICANT: Haataja, Sauli J.
TITLE OF INVENTION: tkla
FILE REFERENCE: GML0210
CURRENT APPLICATION NUMBER: US/09/298,724
CURRENT FILING DATE: 1999-04-23
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1968
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
US-09-298-724-1

Query Match 20.3%; Score 78.2; DB 4; Length 1968;
Best Local Similarity 51.9%; Pred. No. 3.6e-12;
Matches 176; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 17 tgacatcattgggttgaggcagcgctccgagctgagatcgaggcagcgccagga 76
Db 1629 tgatctatctatctatctgctacagatcgagtgatctgagctacacagcgccagga 1688
QY 77 gctgaggaagggaggaagaaagcgctcgctcgctcgctcgctcgctcgctcgctcg 136
Db 1689 attggttttaacaaggcggttaagtcgltggtatctatctgacctcaaccgaacttga 1748
QY 137 tgagcagtcggaatgagtaacaagagcgctccctcgccgagcgcacagcgagatcag 196
Db 1749 tgctcaagatgcttacctacaagaaagacatttaccacctaagactcgctcgltgagc 1808
QY 197 catcgagcgccggttcactctcgctgagcaagaatcagtcgagcccaaggaagggccat 256
Db 1809 catgtaaatgagcagcgcccaagttggtaacaagtattggtttggtgagtcgagcggtcat 1868
QY 257 tgagctcgacaagatcgcgcgagtgctcctcgccgagcgagctctacaagagatcgagat 316
Db 1869 cggtattgacatcttcggtcgctcgctcgcccaagctcagactgagttgataattgagtt 1928
QY 317 cagcgtagagagcatcattgcaactgccaagagcttta 355
Db 1929 taaggtagaataatcggttgctcaagttaaagttacccctata 1967

RESULT 7

PCT-US96-05320A-1025
Sequence 1025, Application PC/TUS9605320A
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences
APPLICANT: 9410 Key West Avenue
APPLICANT: Rockville, MD 20850
APPLICANT: United States of America
APPLICANT: Johns Hopkins University
APPLICANT: 720 Rutland Avenue
APPLICANT: Baltimore, MD 21205

APPLICANT: United States of America
APPLICANT: Mark D. Adams
APPLICANT: Owen White
APPLICANT: Hamilton O. Smith
APPLICANT: J. Craig Venter
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20003-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/05320A
FILING DATE: April 12, 1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: June 7, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: June 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Eric K. Steife
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.014PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 1025:
SEQUENCE CHARACTERISTICS:
LENGTH: 1995 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
PCT-US96-05320A-1025

Query Match 19.9%; Score 76.8; DB 5; Length 1995;
Best Local Similarity 51.8%; Pred. No. 8.5e-12;
Matches 174; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

QY 11 caagcttgacatcattgggttgaggcagcgctccgagctgagatcgaggcaatcgagc 70
Db 1650 CACACCTGAGTTAATCTTCAATTCGACAGAGTTCTGMAAGTGGAATTAGCGTCAAGCAGC 1709
QY 71 cgaagcctgaggaaggaaggaagcggtccggtcgctcgctcgctcgctcgctcgctcg 130
Db 1710 CGAAGCATTAAGTGCCTGAAGGTAAGAAAGTACGCGTGTGTTATGCGCAAGTACTAACCG 1769
QY 131 ctttgatgagcagctcgagatgatacaagagagcgctccctcgccgagcgtaacagcgag 190
Db 1770 TTTGATTAACAAGATGACGCTACCGTGAAGTGTATTACTGACGCGGTATACCAACG 1829
QY 191 gttcagcatcgagcgccggttcactctcggtcgagagaagtaagtcgagcccaagcaaa 250
Db 1830 TGTTCGATTGAAGCGGGATTGCTGACTTCTGTTAATAACGTTGATTCATATGATGTCG 1889
QY 251 ggcattgagcagcaaaagttcgagcgagtgctcctcgccgagcgagatcacaagagta 310
Db 1890 TGTATTGCGGTATGATTAAGTTTGGCGAATCCGACACAGACATCAATTATTCAAACTCTT 1949
QY 311 cggcatcacggtgagagcatcattgcaactgcca 346
Db 1950 TGGTTCACTGTTGAAAACGTGTTAGCAAAAGCGAA 1985

RESULT 8
 US-08-858-207A-35/C
 : Sequence 35, Application US/08858207A
 : Patent No. 6348328
 : GENERAL INFORMATION:
 : APPLICANT: Black, Michael
 : APPLICANT: Hodgson, John
 : APPLICANT: Knowles, David
 : APPLICANT: Nicholas, Richard
 : APPLICANT: Stodola, Robert
 : TITLE OF INVENTION: No. 6348328e1 Compounds
 : NUMBER OF SEQUENCES: 552
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: SmithKline Beecham Corporation
 : STREET: 709 Swedeland Road
 : CITY: King of Prussia
 : STATE: PA
 : COUNTRY: USA
 : ZIP: 19406-0939
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 : SOFTWARE: FastSeq for Windows Version 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/858,207A
 : FILING DATE: 09-MAY-1997
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 60/017670
 : FILING DATE: 14-MAY-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Gimm1, Edward R
 : REGISTRATION NUMBER: 38,891
 : REFERENCE/DOCKET NUMBER: P50475
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 610-270-4478
 : TELEFAX: 610-270-5090
 : TELEX:
 : INFORMATION FOR SEQ ID NO: 35:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1738 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 US-08-858-207A-35

Query Match	16.4%	Score 63	DB 4	Length 1738
Best Local Similarity	51.8%	Pred. No. 4	1e-08	
Matches 141	Conservative	0	Mismatches 131	Indels 0
			Gaps	0
QY	17	tgcacatccttggttggcaccgcgctccgagcttgagatcgcggycaatgcgcgcgaca	76	
Db	272	TGATACTATTTCATTTGCTACAGGATCTCGAGGTCAATCTACTATCAAAAGCTGTCAAGGA	213	
QY	77	gctgagagaagaggggaagacggtctcgcgctgcctctgcgtctctctggaaactttga	136	
Db	212	ATTGGTTTTACAGGTGGTAAAGTACGCTGTGTATCTATGCCCTCAACCGAATATTTGA	153	
QY	137	tgaagcagtcggaatgaatcaagaagagacgctccctccctgcgcgacgctacacgagatcag	196	
Db	152	TGCTCAAGATCTTACTCTTCAAGGAAGACATTTTACCATCTTAAGACTCTGCTGCTGTGGC	93	
QY	197	catcgagcccggtcccaatctccgctcgtgcgacagagtaagctcgcagccaaagaaagccat	256	
Db	92	CATTGAATGCGACGAGCACCACAAAGTTGGTACACAGTATGTGGTTTGGATGGCGCGGNCAAT	33	
QY	257	tggcatcgacaagttccgacgcgcgagtgctctctg	288	
Db	32	CGGTATTGACATCTTCGGTGGCGTCTGCCCCAG	1	

```

1      RESULT          9
2      US-08-743-637B-17
3      : Sequence 17, Application US/08743637B
4      : Patent No. 5994066
5      :
6      : GENERAL INFORMATION:
7      : APPLICANT: BERGERON, Michel G.
8      : APPLICANT: PICARD, Francois J.
9      : APPLICANT: OUELLETTE, Marc
10     : APPLICANT: ROY, Paul H.
11     :
12     : TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
13     : TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
14     : TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
15     : TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...
16     :
17     : NUMBER OF SEQUENCES: 273
18     :
19     : CORRESPONDENCE ADDRESS:
20     : ADDRESSEE: QUARLES & BRADY
21     : STREET: 411 EAST WISCONSIN AVENUE
22     : CITY: MILWAUKEE
23     : STATE: WISCONSIN
24     : COUNTRY: USA
25     :
26     : ZIP: 53202-4497
27     :
28     : COMPUTER READABLE FORM:
29     : MEDIUM TYPE: Floppy disk
30     : COMPUTER: IBM PC compatible
31     : OPERATING SYSTEM: PC-DOS/MS-DOS
32     : SOFTWARE: PatentIn Release #1.0, Version #1.30
33     :
34     : CURRENT APPLICATION DATA:
35     : APPLICATION NUMBER: US/08/743,637B
36     : FILING DATE: 04-NOV-1996
37     :
38     : CLASSIFICATION: 435
39     :
40     : PRIOR APPLICATION DATA:
41     : APPLICATION NUMBER: US 08/526,840
42     : FILING DATE: 11-SEP-1995
43     :
44     : ATTORNEY/AGENT INFORMATION:
45     : NAME: BAKER, Jean C.
46     : REGISTRATION NUMBER: 35,433
47     : REFERENCE/DOCKET NUMBER: 850586.90012
48     :
49     : TELECOMMUNICATION INFORMATION:
50     : TELEPHONE: (414) 277-5000
51     : TELEFAX: (414) 277-5591
52     :
53     : INFORMATION FOR SEQ. ID NO.: 17:
54     : SEQUENCE CHARACTERISTICS:
55     : LENGTH: 1872 base pairs
56     : TYPE: nucleic acid
57     : STRANDEDNESS: double
58     : TOPOLOGY: linear
59     : MOLECULE TYPE: DNA (genomic)
60     : ORIGINAL SOURCE:
61     : ORGANISM: Pseudomonas aeruginosa
62     :
63     : US-08-743-637B-17

```

[illegible]

QY 244 aagcgcaagccattgcatcgacaagaattcgcgcgagtcctctcgccggagcatctaca 303
Db 956 GCGGCGATGACCGTAGACACCCAGACCTACACCCGGAACCGGGCGCTGGCGGCGATCATCG 1015
QY 304 aggaagtcagccatcccgctggagagcatca 333
Db 1016 CCTCCACGACGACCCCGAGTTTCATCGTCA 1045

RESULT 10
US-08-526-840B-17
; Sequence 17, Application US/08526840B
; Patent No. 6001564
; GENERAL INFORMATION:
; APPLICANT: BERGERON, Michel G.
; APPLICANT: OUELLETTE, Marc
; APPLICANT: ROY, Paul H.
; TITLE OF INVENTION: SPECIFIC AND UNIVERSAL PROBES AND
; TITLE OF INVENTION: AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND IDENTIFY
; TITLE OF INVENTION: COMMON BACTERIAL PATHOGENS AND ANTIBIOTIC RESISTANCE
; TITLE OF INVENTION: FROM CLINICAL SPECIMENS FOR ROUTINE DIAGNOSIS IN ...
; NUMBER OF SEQUENCES: 177
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: QUARLES & BRADY
; STREET: 411 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,840B
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/304,732
; FILING DATE: 12-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 850586,90012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5000
; TELEFAX: (414) 277-5591
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1872 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Pseudomonas aeruginosa
US-08-526-840B-17

Query Match 13.4%; Score 51.6; DB 3; Length 1872;
Best Local Similarity 47.3%; Pred. No. 4.8e-05;
Matches 156; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 4 ccggacacaaagcctgacatcatcttggttggcaacggtctccgagcttggagatcgcgga 63
Db 716 CCGCTTCGGCGCCCTACCTGCTGTGTGATCGAGCCGCGACGCGGATCGACATCATCGAC 775
QY 64 atcgcgccgacgaagcttgaaggaaggaagcgatcgcgctgcgtctgcttctct 123
Db 776 TGTGGCCCAAGAGACCCAGCAAGTCCCGAATCAAGATCGGATCGAGGCCGCTCGG 835
QY 124 gggaaaccttggatgagcagtcgagatgatacaagaagagcgctcctcctcgccagctca 183

Db 836 TGGAAAGCTCCAGTTTCAGAGGGCTAGAGAGACCGCTACACATCCCGGCGCTTACTGGC 895
QY 184 cagcgaggtatcagcatcgacgagccggatccactctcgcttgcagaagtcagtcgagccc 243
Db 896 CGCCGAGTTTCGGGATCATGAGACGGGAGACCCCTGGAACCGAAGCATGCTTCACACC 955
QY 244 aagcgcaagccattgcatcgacaagaattcgcgcgagtcctctcgccggagcatctaca 303
Db 956 GCGGCGATGACCGTAGACACCCAGACCTACACCCGGAACCGGGCGCTGGCGGCGATCATCG 1015
QY 304 aggaagtcagccatcccgctggagagcatca 333
Db 1016 CCTCCACGACGACCCCGAGTTTCATCGTCA 1045

RESULT 11
US-09-327-487A-4
; Sequence 4, Application US/09327487A
; Patent No. 6352847
; GENERAL INFORMATION:
; APPLICANT: MATSUKAWA, Hirokazu
; APPLICANT: OKA, Osamu
; APPLICANT: FUJITA, Tutoshi
; APPLICANT: MIYAZAKI, Kentaro
; TITLE OF INVENTION: AMMONIA ELIMINATION LIQUID REAGENT
; FILE REFERENCE: 028022-013
; CURRENT APPLICATION NUMBER: US/09/327,487A
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: JP 10-176643
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1284
; TYPE: DNA
; ORGANISM: Thermus aquaticus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1284)
US-09-327-487A-4

Query Match 11.4%; Score 44; DB 4; Length 1284;
Best Local Similarity 46.4%; Pred. No. 0.0047;
Matches 143; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 3 accggacacaaagcctgacatcatcttggttggcaacggtctccgagcttggagatcgcgga 62
Db 973 atcgcccccgggccacaacatcaacgaagcgggccacgcgctcttggagccacccac 1032
QY 63 aatcgcgccgaagcttgaaggaaggaaggaagcgatcgcgctgcgtctctgctctcc 122
Db 1033 ggcacccgcccccaagtagcttgcgcaagacaagtgaaaccccaagagatcatctctcc 1092
QY 123 tgggaaccttcttgatgagcagatcgatgagtaacaaaggaagcgctcctcctcgccagctc 182
Db 1093 ggggaatgagtcttcgtaaccttggctggaacgagcgcggaacccaatcatcaggggcc 1152
QY 183 aaagcagagatcagatcgagcgcggtccactctcggtctgcggaagtacgtctggagcc 242
Db 1153 atggagagacatcagcaagagggccttgcctcctcaagacttccacgcctcttgglyggcc 1212
QY 243 caagcgaagccatctgcatcgacaagttcggcgcgagttctcctcgcggagacatctac 302
Db 1213 gagggcgaagccgcgcacgctcttctaagacacagcgagttcgcgagcgccctgataccagcac 1272
QY 303 aagagga 310
Db 1273 atggacta 1280

RESULT 12


```

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20..3916
US-08-340-011-1

```

```

Query Match      10.7%; Score 41.2; DB 1; Length 4195;
Best Local Similarity 46.3%; Pred. No. 0.039;
Matches 136; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

```

```

QY 51 gagatcgaggcaatggtggcgagagactgaagaaagggaagacggtccgctgctc 110
   || || || || || || || || || || || || || || || || || || || ||
Db 2075 GACCTCTGTTGACGACGTGAGCGACTCGCTGAGATGCAGTGTGGTGCCGAGCGCAC 2134
QY 111 tcgttcgtctctctgggaactctttgatgagcagtcgagatgatacagaagagcgtcc 170
   || || || || || || || || || || || || || || || || || || || ||
Db 2135 GCGGCCAGCATCGTGTGTACAAAGACGAGAGGCTGCTGAGGAAAAGTGTGAGTCGAC 2194
QY 171 cctgcgagctcacagagagatcagcatcagagccgggtccactctcgctgagagaag 230
   || || || || || || || || || || || || || || || || || || || ||
Db 2195 TTGGCGGACTTCACACGAGAGCTGAGCATCCAGCGCTGCCGAGAGAGATGCGGAGCGC 2254
QY 231 taagtcgagagcccaaggcaagccatgagcatcgacagaagtcggcgaggtgctctgc 290
   || || || || || || || || || || || || || || || || || || || ||
Db 2255 TATCTGTGACAGCGTGTGACACGCGCAAGGCGTGGTCACTCTCCGCCAGCGTGGCGTG 2314
QY 291 gggagcagatctacaagagagatcagcatcaccgtggagagacatctgcaactgcc 344
   || || || || || || || || || || || || || || || || || || || ||
Db 2315 GAAGGCTCCGAGGATTAAGGCGACATGAGATGATGATCTGTGTGCGTACCGGC 2368

```

RESULT 15

```

US-08-901-710-1
; Sequence 1, Application US/08901710
; Patent No. 6107046

```

GENERAL INFORMATION:

```

; APPLICANT: Alltalo, Kari
; APPLICANT: Aprelikova, Olga
; APPLICANT: Pajusola, Katri
; APPLICANT: Armstrong, Elin
; APPLICANT: Korhonen, Jaana
; APPLICANT: Kaipainen, Arja
; APPLICANT: Matikainen, Marja-Terttu
; TITLE OF INVENTION: FLT4, A RECEPTOR TYROSINE KINASE, AND USES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402

```

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/901,710
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,011
; FILING DATE: 14-NOV-1994
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: 08/257,754
; FILING DATE: 09-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/959,951
; FILING DATE: 09-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28113/33824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20..3916
US-08-901-710-1

```

```

Query Match      10.7%; Score 41.2; DB 3; Length 4195;
Best Local Similarity 46.3%; Pred. No. 0.039;
Matches 136; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

```

```

QY 51 gagatcgaggcaatggtggcgagagactgaagaaagggaagacggtccgctgctc 110
   || || || || || || || || || || || || || || || || || || || ||
Db 2075 GACCTCTGTTGACGACGTGAGCGACTCGCTGAGATGCAGTGTGGTGCCGAGCGCAC 2134
QY 111 tcgttcgtctctctgggaactctttgatgagcagtcgagatgatacagaagagcgtcc 170
   || || || || || || || || || || || || || || || || || || || ||
Db 2135 GCGGCCAGCATCGTGTGTACAAAGACGAGAGGCTGCTGAGGAAAAGTGTGAGTCGAC 2194
QY 171 cctgcgagctcacagagagatcagcatcagagccgggtccactctcgctgagagaag 230
   || || || || || || || || || || || || || || || || || || || ||
Db 2195 TTGGCGGACTTCACACGAGAGCTGAGCATCCAGCGCTGCCGAGAGAGATGCGGAGCGC 2254
QY 231 taagtcgagagcccaaggcaagccatgagcatcgacagaagtcggcgaggtgctctgc 290
   || || || || || || || || || || || || || || || || || || || ||
Db 2255 TATCTGTGACAGCGTGTGACACGCGCAAGGCGTGGTCACTCTCCGCCAGCGTGGCGTG 2314
QY 291 gggagcagatctacaagagagatcagcatcaccgtggagagacatctgcaactgcc 344
   || || || || || || || || || || || || || || || || || || || ||
Db 2315 GAAGGCTCCGAGGATTAAGGCGACATGAGATGATGATCTGTGTGCGTACCGGC 2368

```

```

Search completed: July 3, 2002, 08:11:38
Job time: 25863 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 05:14:46 ; Search time 5254.53 Seconds

(without alignments)
1051.399 Million cell updates/sec

Title: US-09-300-482-569

Perfect score: 264
Sequence: 1 ctgactgcagcaattccca.....gacatgcttcagatctta 264

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
------------	-------------	--------------	-------	-------------

1	51.2	19.4	7218	6	166494	166494 Sequence 14
2	41.8	15.8	197284	9	CNS01DWL	AL137230 Human chr
3	38.4	14.5	156062	2	AC097255	AC097255 Rattus no
4	38.2	14.5	1552	8	STU95923	U95923 Solanum tub
5	38.2	14.5	15313	2	AC101327	AC101327 Mus muscu
6	38.2	14.5	146741	2	AC011452	AC011452 Homo sapi
7	38.2	14.5	168230	2	AC010624	AC010624 Homo sapi
8	38.2	14.5	175999	2	AC021154	AC021154 Homo sapi
9	38.2	14.4	175602	2	AC106542	AC106542 Rattus no
10	37.6	14.2	167254	2	CNS05TDS	AL357093 Human chr
11	37.6	14.2	167390	2	CNS07263	AC007263 Homo sapi
12	37.6	14.2	200542	2	CNS01R83	AL157736 Homo sapi
13	37.6	14.2	233528	2	AC099382	AC099382 Rattus no
14	37.4	14.2	139005	2	AF286112	AF286112 Homo sapi
15	37.4	14.2	165901	2	AF003474	AF003474 Homo sapi
16	37.4	14.2	171427	9	AC021590	AC021590 Homo sapi
17	37.2	14.1	16930	2	AF294352	AF294352 Homo sapi
18	37.2	14.1	79452	2	AC016536	AC016536 Homo sapi
19	37.2	14.1	107109	2	AC01781B1	AL118522 Human DNA
20	37.2	14.0	158641	2	AC098454	AC098454 Rattus no
21	37.2	14.0	179604	2	AP002959	AP002959 Homo sapi
22	36.8	13.9	73625	2	AC108383	AC108383 Pan trogl
23	36.8	13.9	139273	2	AL596275	AL596275 Human DNA
24	36.8	13.9	187999	2	AC026928	AC026928 Homo sapi
25	36.8	13.9	205504	2	AC022699	AC022699 Mus muscu
26	36.6	13.9	24721	2	HS1190H5	Z68163 Human DNA s
27	36.6	13.9	156727	2	AL627084	AL627084 Homo sapi
28	36.6	13.8	200622	2	AL158068	AL158068 Homo sapi
29	36.4	13.8	54929	2	AC105158	AC105158 Mus muscu
30	36.4	13.8	144107	2	AC012027	AC012027 Homo sapi
31	36.4	13.8	147610	2	AC099116	AC099116 Rattus no
32	36.4	13.8	170048	2	AC089988	AC089988 Homo sapi
33	36.4	13.8	192182	2	AC094597	AC094597 Rattus no
34	36.4	13.8	195052	2	AC106707	AC106707 Homo sapi
35	36.4	13.8	212382	2	AC010936	AC010936 Homo sapi
36	36.4	13.8	215196	2	AL603705	AL603705 Mus muscu
37	36.2	13.7	28812	2	AC102623	AC102623 Mus muscu
38	36.2	13.7	62109	9	AL359984	AL359984 Human DNA
39	36.2	13.6	569	11	HUMUT5530	L30662 Human STS U
40	36.2	13.6	111845	2	AC097570	AC097570 Rattus no
41	36.2	13.6	152366	9	AL138894	AL138894 Human DNA
42	36.2	13.6	184092	2	AC084010	AC084010 Homo sapi
43	36.2	13.6	200925	2	AC009090	AC009090 Homo sapi
44	36.2	13.6	298166	2	AC087563	AC087563 Homo sapi
45	35.8	13.6	180261	2	AC090308	AC090308 Homo sapi

ALIGNMENTS

RESULT	1	LOCUS	166494	SEQUENCE	14 from patent US 5670367.	DNA	linear	PAT	28-DEC-1997
DEFINITION	166494	Sequence 14 from patent US 5670367.							
ACCESSION	166494	GI:2724471							
VERSION	166494.1	GI:2724471							
KEYWORDS	Unknown.								
SOURCE	Unknown.								
ORGANISM	Unknown.								
REFERENCE	1 (bases 1 to 7218)								
AUTHORS	Dorner, F., Scheiflinger, F. and Falkner, F. Gunter.								
TITLE	Recombinant fowlpox virus								
JOURNAL	Patent: US 5670367-A 14 23-SEP-1997;								
FEATURES	Location/Qualifiers								
source	1..7218								
BASE COUNT	1944 a 1491 c 1486 g 1929 t 368 others								
ORIGIN	/organism="unknown"								

Query Match 19.4%; Score 51.2; DB 6; Length 7218;
Best Local Similarity 5.5%; Pred. No. 0.00053;

Louisegeed, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M.,
Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, J., Mohabbat, K.,
Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nkokenko, S.,
Ogun, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M.,
Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoostari, N.,
Slason, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H.,
Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,
Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
Washington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S.,
Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 156062)
Morley, K.C.

Submitted (13-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:17062557.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center Project name: GIRI
Center Clone name: CH230-152G23
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 109492 bases at least Q40
Consensus quality: 119842 bases at least Q30
Consensus quality: 128566 bases at least Q20
Estimated insert size: 114589; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 1.4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 67 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 6695: contig of 6694 bp in length
* 6795: gap of unknown length
* 13119: contig of 6324 bp in length
* 13219: gap of unknown length
* 16527: contig of 3309 bp in length
* 16628: gap of unknown length
* 21815: contig of 5188 bp in length
* 21816: gap of unknown length
* 21916: contig of 4942 bp in length
* 26858: gap of unknown length
* 26958: gap of unknown length
* 31437: contig of 4479 bp in length
* 31537: gap of unknown length
* 35430: contig of 3894 bp in length
* 35530: gap of unknown length
* 37687: contig of 2157 bp in length
* 37688: gap of unknown length
* 41604: contig of 3817 bp in length
* 41605: gap of unknown length

41705: contig of 3140 bp in length
44845: gap of unknown length
44945: contig of 3191 bp in length
48136: gap of unknown length
48236: contig of 3025 bp in length
51261: gap of unknown length
51361: contig of 2890 bp in length
54251: gap of unknown length
54350: contig of 3005 bp in length
57356: gap of unknown length
57455: contig of 2030 bp in length
57456: gap of unknown length
59485: contig of 2725 bp in length
59486: gap of unknown length
59586: contig of 2725 bp in length
62310: gap of unknown length
62411: contig of 3086 bp in length
65497: gap of unknown length
65597: contig of 2655 bp in length
68252: gap of unknown length
68351: contig of 3140 bp in length
68352: gap of unknown length
71491: contig of 1110 bp in length
71492: gap of unknown length
71592: contig of 1110 bp in length
72702: gap of unknown length
72801: contig of 1888 bp in length
74680: gap of unknown length
74789: contig of 2674 bp in length
77463: gap of unknown length
77464: contig of 2275 bp in length
77564: gap of unknown length
79838: contig of 2518 bp in length
79839: gap of unknown length
82457: contig of 2069 bp in length
82556: gap of unknown length
84625: contig of 1389 bp in length
84626: gap of unknown length
84726: contig of 3020 bp in length
86115: gap of unknown length
86214: contig of 2614 bp in length
89235: gap of unknown length
89236: contig of 2614 bp in length
89335: gap of unknown length
91948: contig of 2161 bp in length
91949: gap of unknown length
92048: contig of 1592 bp in length
94209: gap of unknown length
94309: contig of 1709 bp in length
94310: gap of unknown length
95901: contig of 1787 bp in length
95902: gap of unknown length
96001: contig of 2304 bp in length
97710: gap of unknown length
97711: gap of unknown length
97810: contig of 2236 bp in length
99597: gap of unknown length
99598: contig of 1981 bp in length
99698: gap of unknown length
102001: contig of 1823 bp in length
102002: gap of unknown length
102101: contig of 1110 bp in length
102102: gap of unknown length
104337: contig of 1294 bp in length
104338: gap of unknown length
104437: contig of 1823 bp in length
105785: gap of unknown length
105884: contig of 1294 bp in length
105885: gap of unknown length
107826: contig of 1823 bp in length
107827: gap of unknown length
107925: contig of 1219 bp in length
107926: gap of unknown length
109144: contig of 1981 bp in length
109145: gap of unknown length
109244: contig of 2243 bp in length
109245: gap of unknown length
111225: contig of 1634 bp in length
111226: gap of unknown length
111326: contig of 2007 bp in length
111327: gap of unknown length
113369: contig of 1634 bp in length
113370: gap of unknown length
113659: contig of 1634 bp in length
113660: gap of unknown length
115302: contig of 1634 bp in length
115303: gap of unknown length
115403: contig of 2007 bp in length
115404: gap of unknown length
117410: contig of 1086 bp in length
117411: gap of unknown length
117510: contig of 1094 bp in length
117511: gap of unknown length
118595: contig of 1094 bp in length
118596: gap of unknown length
118686: contig of 1294 bp in length
118687: gap of unknown length
119790: contig of 1294 bp in length
119791: gap of unknown length
121184: contig of 1823 bp in length
121185: gap of unknown length
123107: contig of 1823 bp in length
123108: gap of unknown length
124316: contig of 1110 bp in length

Center: Whitehead Institute/ MIT Center for Genome Research


```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0510116
----- Summary Statistics -----
Sequencing vector: plasmid; 22%
Chemistry: Dye-primer ET; 78% of reads
Chemistry: Dye-terminator Big Dye; 22% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 165027 bases at least Q40
Consensus quality: 168606 bases at least Q30
Consensus quality: 170831 bases at least Q20
Insert size: 18800; agarose-fp
Insert size: 175690; sum-of-ctrls
Quality coverage: 3.68 in Q20 bases; agarose-fp
Quality coverage: 3.97 in Q20 bases; sum-of-ctrls
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
* 1820: contig of 1820 bp in length
*
* 1821 1920: gap of unknown length
*
* 1921 4432: contig of 2512 bp in length
*
* 4433 4532: gap of unknown length
*
* 4533 6592: contig of 2460 bp in length
*
* 6593 7092: gap of unknown length
*
* 7093 10867: contig of 3775 bp in length
*
* 10868 10967: gap of unknown length
*
* 10968 14067: contig of 3100 bp in length
*
* 14068 14167: gap of unknown length
*
* 14168 18869: contig of 4702 bp in length
*
* 18870 18969: gap of unknown length
*
* 18970 24630: contig of 5661 bp in length

```


*	24631	24730:	gap of unknown length
*	24711	31836:	contig of 7106 bp in length
*	31897	31936:	gap of unknown length
*	31987	38296:	contig of 7360 bp in length
*	39287	39396:	gap of unknown length
*	39337	45890:	contig of 6494 bp in length
*	45891	45990:	gap of unknown length
*	45921	56224:	contig of 10234 bp in length
*	56225	56324:	gap of unknown length
*	64405	64404:	contig of 8080 bp in length
*	64405	72478:	gap of unknown length
*	72479	72578:	gap of unknown length
*	72579	79959:	contig of 7381 bp in length
*	79960	80059:	gap of unknown length
*	80060	91688:	contig of 11629 bp in length
*	91689	91788:	gap of unknown length
*	91789	104423:	contig of 12635 bp in length
*	104424	104523:	gap of unknown length
*	104524	123430:	contig of 18907 bp in length
*	123431	123530:	gap of unknown length
*	12351	149278:	contig of 28748 bp in length
*	149279	149378:	gap of unknown length
*	149379	175999:	contig of 26621 bp in length

source	1. .175999
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="19"
	/clone="RP11-510116"
misc_feature	1. .1820
	/note="assembly_name:Contig6"
misc_feature	1921. .4432
	/note="assembly_name:Contig7"
misc_feature	4533. .6992
	/note="assembly_name:Contig8"
misc_feature	7093. .10867
	/note="assembly_name:Contig9"
misc_feature	10968. .11067
	/note="assembly_name:Contig10"
misc_feature	14168. .18869
	/note="assembly_name:Contig11"
misc_feature	18970. .24630
	/note="assembly_name:Contig12"
misc_feature	24731. .31836
	/note="assembly_name:Contig13"
misc_feature	31937. .39296
	/note="assembly_name:Contig14"
misc_feature	39397. .45890
	/note="assembly_name:Contig15"
misc_feature	45991. .56224
	/note="assembly_name:Contig16"
misc_feature	56325. .64404
	/note="assembly_name:Contig17"
misc_feature	64505. .72478
	/note="assembly_name:Contig18"
misc_feature	72579. .79959
	/note="assembly_name:Contig19"
misc_feature	80060. .91688
	/note="assembly_name:Contig20"
misc_feature	91789. .104423
	/note="assembly_name:Contig21"
misc_feature	104524. .123430
	/note="assembly_name:Contig22"
misc_feature	123531. .149278
	/note="assembly_name:Contig23"
misc_feature	149379. .175999

[illegible]

Query Match	14.5%	Score 38.2;	DB 2;	Length 175999;
Best Local Similarity	61.6%;	Pred. No. 1.3;		
Matches 61; Conservative	0;	Mismatches 38;	Indels 0;	Gaps 0;

[illegible]

LOCUS	DEFINITION	SEQUENCING IN PROGRESS	HTG 12-JAN-2002
AC106542	Rattus norvegicus clone CH230-91p1, 175602 bp	***	linear

ACCESSION	AC106542
VERSION	AC106542.1
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	Norway rat.
ORGANISM	Rattus norvegicus

REFERENCE
AUTHORS

1 (bases 1 to 175602)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,

Benton, J., Blimie, R., Blankenburg, K., Bonnin, D., Bouck, J., Bowler, S., Brieva, M., Brown, E., Brown, M., Bryant, N. P., Bulhary, C., Burck, P., Burlett, C., Burrell, K. L., Byrd, N. C., Carion, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D., Chen, G., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R., Davila, M. L., Davis, C., Davy-Carroll, L., Dederich, D. A., Delaney, K. R., Delgado, O., Denn, A. L., Ding, Y., Dinh, H. H., Douthwaite, K. J., Draper, H., Dugan-Rocha, S., Durbin, K. J., Eamhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulyk, S., Hunne, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudash, S., Karlsson, E., Kelly, S., Khan, U., Kling, L., Korah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L., Li, Z., Li, Z., Lichaarge, O., Lien, C., Liu, Z., Liu, W., Louisaged, H., Lozada, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., Mcleod, M. P., Meador, M., Mel, G., Metzker, M., Miner, G., Miner, Z., Mitchell, J., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkwo, S., Ogulu, M., Okwuonon, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L., Ruiles, M., Ren, Y., Rivers, M., Rojas, A., Rojoubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, L., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Swalek, A., Tabori, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlaczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D., Weinrock, G., and Gibbs, R.

TITLE	Direct Submission
JOURNAL	unpublished
REFERENCE	2 (bases 1 to 175602)
AUTHORS	Morley, R.C.
TITLE	Direct Submission
JOURNAL	Submitted (12-JAN-2002) Human Genome Sequencing Center, Department

COMMENT

of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GLTB

Center clone name: CH230-91P1

----- Summary Statistics

Assembly program: Phrap; version 0.990329First call to
findHirepList

Consensus quality: 150917 bases at least Q40

Consensus quality: 158026 bases at least Q30

Consensus quality: 164737 bases at least Q20

Estimated insert size: 150665; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-IP estimation

Quality coverage: 2.4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 61 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

*-----

* 1 8244: contig of 8244 bp in length

* 8245 8344: gap of unknown length

* 8345 14013: contig of 3669 bp in length

* 14014 14113: gap of unknown length

* 14114 18063: contig of 3950 bp in length

* 18064 18163: gap of unknown length

* 18164 22920: contig of 4757 bp in length

* 22921 23020: gap of unknown length

* 23021 27998: contig of 4978 bp in length

* 27999 28098: gap of unknown length

* 28099 34626: contig of 6528 bp in length

* 34627 34726: gap of unknown length

* 34727 41062: contig of 6336 bp in length

* 41063 41162: gap of unknown length

* 41163 45112: contig of 3950 bp in length

* 45113 45212: gap of unknown length

* 45213 50721: contig of 5509 bp in length

* 50722 50821: gap of unknown length

* 50822 55943: contig of 5122 bp in length

* 55944 56043: gap of unknown length

* 56044 59414: contig of 3371 bp in length

* 59415 59514: gap of unknown length

* 59515 62979: contig of 3465 bp in length

* 62980 63079: gap of unknown length

* 63080 67194: contig of 4115 bp in length

* 67195 67294: gap of unknown length

* 67295 70427: contig of 3133 bp in length

* 70428 70527: gap of unknown length

* 70528 73739: contig of 3212 bp in length

* 73740 73839: gap of unknown length

* 73840 78179: contig of 4340 bp in length

* 78180 78279: gap of unknown length

* 78280 82269: contig of 3990 bp in length

* 82270 82369: gap of unknown length

* 82370 87355: contig of 5186 bp in length

* 87356 87655: gap of unknown length

* 87656 91236: contig of 3581 bp in length

* 91237 91336: gap of unknown length

* 91337 95629: contig of 4193 bp in length

* 95630 95530: gap of unknown length

* 95531 97911: contig of 2282 bp in length

* 97912 98011: gap of unknown length

* 98012 101900: contig of 3889 bp in length

* 101901 102000: gap of unknown length

* 102001 105753: contig of 3753 bp in length

* 105754 105853: gap of unknown length

* 105854 109202: contig of 3349 bp in length

* 109203 109302: gap of unknown length

* 109303 112783: contig of 3481 bp in length

* 112784 112883: gap of unknown length

* 112884 116644: contig of 3761 bp in length

* 116645 116744: gap of unknown length

* 116745 118406: contig of 1662 bp in length

* 118407 118506: gap of unknown length

* 118507 121212: contig of 2706 bp in length

* 121213 121312: gap of unknown length

* 121313 124459: contig of 3147 bp in length

* 124460 124559: gap of unknown length

* 124560 126607: contig of 2048 bp in length

* 126608 126707: gap of unknown length

* 126708 128164: contig of 2457 bp in length

* 128165 129165: gap of unknown length

* 129166 129264: gap of unknown length

* 129265 131628: contig of 2364 bp in length

* 131629 131728: gap of unknown length

* 131729 133292: contig of 1564 bp in length

* 133293 133392: gap of unknown length

* 133393 134786: contig of 1394 bp in length

* 134787 134886: gap of unknown length

* 134887 136549: contig of 1663 bp in length

* 136550 136650: gap of unknown length

* 136651 138272: contig of 1623 bp in length

* 138273 138372: gap of unknown length

* 138373 139458: contig of 1086 bp in length

* 139459 139558: gap of unknown length

* 139559 141659: gap of unknown length

* 141660 141759: contig of 2101 bp in length

* 141760 143776: gap of unknown length

* 143777 143876: contig of 2017 bp in length

* 143877 144923: gap of unknown length

* 144924 145023: contig of 1047 bp in length

* 145024 146165: gap of unknown length

* 146166 146265: contig of 1142 bp in length

* 146266 147462: gap of unknown length

* 147463 147562: contig of 1197 bp in length

* 147563 149380: gap of unknown length

* 149381 149480: contig of 1818 bp in length

* 149481 150621: gap of unknown length

* 150622 150721: contig of 1141 bp in length

* 150722 152401: gap of unknown length

* 152402 152501: contig of 1680 bp in length

* 152502 154239: gap of unknown length

* 154240 154339: contig of 1738 bp in length

* 154340 155389: gap of unknown length

* 155390 155688: contig of 1249 bp in length

* 155689 155801: gap of unknown length

* 155802 156801: contig of 1113 bp in length

* 156802 156901: gap of unknown length

* 156902 158250: contig of 1349 bp in length

* 158251 158350: gap of unknown length

* 158351 159522: contig of 1172 bp in length

* 159523 159622: gap of unknown length

* 159623 160717: contig of 1095 bp in length

* 160718 160817: gap of unknown length

* 160818 162530: contig of 1713 bp in length

* 162531 162630: gap of unknown length

* 162631 164499: contig of 1869 bp in length

* 164500 164599: gap of unknown length

* 164600 166124: contig of 1525 bp in length

* 166125 166224: gap of unknown length

* 166225 167248: contig of 1024 bp in length

* 167249 167348: gap of unknown length

* 167349 168555: contig of 1307 bp in length

* 168556 168755: gap of unknown length

* 168756 169814: contig of 1059 bp in length

Query Match
Best Local Similarity
14.4%: Score 38; DB 2; Length 175602;
53.3%: Pred. No. 1.5;

Matches	80;	Conservative	0;	Mismatches	70;	Indels	0;	Gaps	0;
QY	49	caaatcacgtctgcttcctatataagaagcgctgcgcccgagattcgtcttccact	108						
Db	71695	CATCTCCCTTCCTCCTTCCGTGGACATAGCCCCCTCTCTCCACTCWCACCCTCCCTCTCT	71636						
OY	109	tcaaacctctttcatcgccttttaatacaaaacctaactatcgcgtctccctcgatcca	168						
Db	71635	TTTTCCCTCCCTCTTCTCCCTCTTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	71576						
QY	169	ttcacctgcgtccctctcttaaaaacctcttac	198						
Db	71575	TTCCCTCCCTCCCT	71546						
RESULT	10								
CNS05TDS		CNS05TDS	167254 bp DNA	Linear	PRI 19-DEC-2001				
LOCUS		Human chromosome 14 DNA sequence BAC R-356K23 of library RPEI-11							
DEFINITION		from chromosome 14 of Homo sapiens (Human), complete sequence.							
ACCESSION		AL357093							
VERSION		AL357093.4 GI:17974147							
KEYWORDS		HTG.							
SOURCE		human.							
ORGANISM		Homo sapiens							
REFERENCE		Eukariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 167254)							
AUTHORS		Hellä,R., Pettit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P., Brother,P., Catiolico,L., Barbe,V., Pelletier,E., Artiguenave,F., Ley,M., Eckenberg,R., Bruls,T., debernardinis,V., Cruaud,C., Guyay,G., Saurin,W. and Weissenbach,J. Sequencing of the human chromosome 14 Unpublished 2 (bases 1 to 167254) Genoscope.							
TITLE		Direct Submission							
JOURNAL		Submitted (19-DEC-2001) Genoscope - Centre National de Sequencage :							
REFERENCE		Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr							
AUTHORS		- Web : www.genoscope.cns.fr)							
TITLE		On Dec 20, 2001 this sequence version replaced gi:15282117.							
JOURNAL		----- Genome Center							
COMMENT		Center: Genoscope / Centre National de Sequencage							
		Center code: GS							
		Web site: http://www.genoscope.cns.fr/							
		Contact: SeqRef@genoscope.cns.fr							

		The following BAC sequence is oriented from the T7 to the SP6 end.							
		Upstream BAC (overlapping the T7 end) : R-945C20 (AC=AL157736)							
		Downstream BAC (overlapping the SP6 end) : R-681H18 -----							
		Summary Statistics							
		Assembly program: Phrap; version 2.0							
		Quality coverage: 9.90x in Q20 bases; sum-of-contigs							

		Overall quality chart :							
		Range : bases							
		0 :							
		1 - 9 :							
		10 - 19 :	1						
		20 - 29 :	10						
		30 - 39 :	432						
		40 - 49 :	1125						
		50 - 59 :	3405						
		60 - 69 :	6191						
		70 - 79 :	17130						
		80 - 89 :	46329						
		90 - 99 :	92631						
		Percentage of bases with a quality value >= 40 : 99 %.							
		Location/Qualifiers							
		I..167254							
		/organism="Homo sapiens"							
FEATURES									
SOURCE									

[illegible]


```

REFERENCE
AUTHORS      2 (bases 1 to 167390)
              Rowen,L., Madan,A., Qln,S., Abbasi,N., Dots,M., Dickhoff,R.,
              James,R., Loretz,C., Lasky,S., Madan,A., Prescott,S., Ratcliffe,A.,
              Shaffer,T. and Hood,L.
TITLE        Direct Submission
JOURNAL      Submitted (08-APR-1999) Multimegabase Sequencing Center, University
              of Washington, PO BOX 357730, Seattle, WA 98195, USA
REFERENCE    3 (bases 1 to 167390)
AUTHORS      Rowen,L., Madan,A., Qln,S., Abbasi,N., Baradaran,L., Birditt,B.,
              Bloom,S., Dots,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
              James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
              and Hood,L.
TITLE        Direct Submission
JOURNAL      Submitted (15-DEC-1999) Multimegabase Sequencing Center, University
              of Washington, PO BOX 357730, Seattle, WA 98195, USA
COMMENT      On Dec 15, 1999 this sequence version replaced gi:4885690.
              -----
              Genome Center
              Center: Multimegabase Sequencing Center
              Center code: UMWSC
              Web site: http://chroma.mbt.washington.edu/msg_www
              Contact: leerowen@u.washington.edu
              -----
              Summary Statistics
              Sequencing vector: pUC18: 108752
              Chemistry: Dye-terminator Big Dye; 90% of reads
              Chemistry: Dye-primer Big Dye; 10% of reads
              Assembly program: Phrap; Version 0.990399
              -----
              Location/Qualifiers
              1. 167390
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="14"
                  /map="14q31"
                  /clone="RP11-79J20"
                  /clone_lib="RP11 human BAC library 11"
                  join(31458..31784,33146..33433)
                  /note="Exact match to cDNA in AF088054 and several ESTs.
                  However there is not a consensus splice site and the
                  sequence in this region is of high quality"
                  32010..32190
                  /note="Low quality data"
                  33800..34070
                  /note="Low quality data -- high GC and compressed region
                  that was resistant to several finishing strategies"
                  33915..>99451
                  /gene="CHES1"
                  /note="Checkpoint suppressor 1: Matches U68723 and several
                  ESTs. The 5' UTR goes from 33915-34074 based on EST
                  AI393423. The 5'UTR of U68723 is not in this sequence.
                  Exon 2 starts at 38455"
                  join(38469..39011,99315..>99451)
                  /gene="CHES1"
                  /codon_start=1
                  /product="checkpoint suppressor 1"
                  /protein_id="AAPI8259.1"
                  /db_xref="GI:6579249"
                  /translation="MGFWMPKKPPSSGISVSGLSQCYGGSGFSKALQEDDDLDS
                  LPDRLREGAMDEDELTNLWMLHESKNLKSFGESYLSVSPVDLDDTPSPAHSD
                  MPIDARPNCKRPYFSCSLIFMALIEDSPKRLPKVDLYNMILHEFPVENAPGKMN
                  SVRNLSLINCFFKVKDKERSOSIGKSLICDPEYRONLDALKKTPYHPHFVFNTP
                  PTCPOAYO"
                  89450..89470
                  /gene="CHES1"
                  /note="Low quality data"
BASE COUNT   42207 a 36473 c 39730 g 48980 t
ORIGIN
Query Match      14.2%: Score 37.6; DB 9; Length 167390;
Best Local Similarity 59.3%; Pred. No. 1.g;
Matches 64; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
dy 76 cgcctgcgcccgagattctgcctctccatcaacacctctccatcgctttaatc 135

```

```

Db      89650   CTGCTTCCCTCCTCCCTCCTCATCTTCCTTCCTCCCTCCCTTCCTCCCTCTCTTTTACC 89709
Oy      136     acaaacataccatcgcgttcctccagacatcatcactgctccctc 183
Db      89710   TCTCTCCTGTCTCTCCTGCCTCCCTCCCTTCCTTCCTTCCTTCCTC 89757

RESULT  12
LOCUS   CNS01RG3                                200542 bp    DNA        linear   HTG 20-SEP-2000
DEFINITION Homo sapiens chromosome 14 clone R-945C20, *** SEQUENCING IN PROGRESS ***, 2 ordered pieces.
ACCESSION AL157736
VERSION   AL157736.3 GI:10279652
KEYWORDS  HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 200542)
AUTHORS   Genoscope.
JOURNAL   Direct Submission
           Submitted (20-SEP-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
COMMENT   On Sep 22, 2000 this sequence version replaced gi:9212407.
           ----- Genome Center
           Center: Genoscope / Centre National de Sequencage
           Center code: GS
           Web site: http://www.genoscope.cns.fr/
           Contact: SeqRef@genoscope.cns.fr
           -----
IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continue. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, phage, etc. . . even if efforts are made to eliminate these contaminating sequences. The following BAC sequence is oriented from the T7 to the Sp6 end.
Upstream BAC (overlapping the T7 end) : R-33N16
Downstream BAC (overlapping the Sp6 end) : R-356K23 Contigs composition :
97511 bp contig from 1 to 97511
102931 bp contig from 97612 to 200542
-----
Overall quality chart :
Range       : bases
0            : 141
1 - 9       : 1506
10 - 19    : 4137
20 - 29    : 9801
30 - 39    : 20690
40 - 49    : 19461
50 - 59    : 27004
60 - 69    : 45678
70 - 79    : 47915
80 - 89    : 18485
90 - 99    : 5724
-----
Percentage of bases with a quality value >= 40 : 81 %.
* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
  1 97511: contig of 97511 bp in length
* 97512 97611: gap of 100 bp

```



```

FEATURES      * 97612 200542: contig of 102931 bp in length.
Source
  Location/Qualifiers
    1. 200542
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /chromosome="14"
       /clone="R-945C20"
       /clone_1lb="RPCI-11"
       122683..122976
       /note="matching EMBL:Z50978"
STS
  RHdb:RH85673
  RHdb:RH42088
  RHdb:RH31324
  dbSTS:STS12299
  Identified using the e-PCR software (G. Schuler)"
  149549..149727
  /note="matching EMBL:R01198"
STS
  RHdb:RH99189
  dbSTS:STS68937
  Identified using the e-PCR software (G. Schuler)"
  189120..189301
  /note="matching EMBL:G33025"
  RHdb:RH67719
  dbSTS:STS47646
  Identified using the e-PCR software (G. Schuler)"
BASE COUNT    50658 a 43824 c 46992 g 58964 t 104 others
ORIGIN
Query Match    14.2% Score 37.6; DB 2; Length 200542;
Best Local Similarity 59.3% Pred. No. 1.8;
Matches 64; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 76 cgcctgcggccgcgagatctcgcctcctcctcaaccccttcacgcgtttatc 135
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 98431 CTCCTTCCCTCCCTCCCTCCCTCACTCTTCCCTCCCTCCCTCCCTCTTTTAC 98490
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 136 acaaaactaactatcgcgttcgcgtatcgcgtatcgcgttcgcgttc 183
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 98491 TCCTCCCTGCTCTGCTGCTCCCTCCCTCCCTCTTCCCTTCTCCCTC 98538
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
AC099382      233528 bp DNA linear HTG 21-DEC-2001
LOCUS        Rattus norvegicus chromosome Msc1 clone CH230-69114, WORKING DRAFT
DEFINITION   AC099382
ACCESSION    AC099382
VERSION      AC099382.3 GI:17974805
KEYWORDS     HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE       Norway rat.
ORGANISM     Rattus norvegicus
              Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murine;
              Rattus
              1 (bases 1 to 233528)
              Wuzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
              Alsbrooks,S.L., Amaralunge,H.C., Aze,J.R., Banks,T., Barbata,J.,
              Benton,J., Bimega,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
              Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
              Burck,P., Burckett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
              Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
              Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
              Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
              Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
              Dean,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
              Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
              Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,D.,
              Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
              Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
              Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
              Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
              Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Humme,J.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Koryah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W.,
Lousaged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mel,G., Metzger,M., Miner,Z., Mitchell,T., Mohabhat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nnokenkwo,S.,
Ogun,M., Okwona,G., Oragunye,N., Oyiedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojuboaka,I., Rolfe,M.,
Rutz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshitari,N.,
Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Swatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Verra,V., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 233528)
Worley,K.C.
Direct Submission
Submitted (10-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 21, 2001 this sequence version replaced gi:17062954.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc@bcm.tmc.edu
----- Project Information
Center project name: GJTF
Center clone name: CH230-69114
----- Summary Statistics
Assembly program: Phrap; version 0.990329first call to
findPhrapList
Consensus quality: 222841 bases at least Q40
Consensus quality: 224953 bases at least Q30
Consensus quality: 226608 bases at least Q20
Estimated insert size: 224352; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agnrose-fp estimation
Quality coverage: 5.7x in Q20 bases; sum-of-contigs estimation
-----
** NOTE: Estimated insert size may differ from sequence length
** (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
** NOTE: This is a 'working draft' sequence. It currently
** consists of 20 contigs. The true order of the pieces
** is not known and their order in this sequence record is
** arbitrary. Gaps between the contigs are represented as
** runs of N, but the exact sizes of the gaps are unknown.
** This record will be updated with the finished sequence
** as soon as it is available and the accession number will
** be preserved.
* 1
* 27123: contig of 27123 bp in length
* 27124 27223: gap of unknown length
* 27224 56657: contig of 29434 bp in length
* 56658 56757: gap of unknown length
* 56758 86690: contig of 29933 bp in length
* 86691 86790: gap of unknown length
* 86791 109359: contig of 22569 bp in length
* 109360 109459: gap of unknown length
* 109460 126882: contig of 17423 bp in length
* 126883 126982: gap of unknown length
* 126983 142143: contig of 15160 bp in length
* 142143 142242: gap of unknown length
* 142243 154308: contig of 12066 bp in length
* 154309 154408: gap of unknown length

```


	*	154409	165642:	contig of 11234 bp in length
	*	165643	165742:	gap of unknown length
	*	165743	176860:	contig of 11118 bp in length
	*	176861	176960:	gap of unknown length
	*	176961	187401:	contig of 10441 bp in length
	*	187402	187501:	gap of unknown length
	*	187502	196368:	contig of 8867 bp in length
	*	196369	196468:	gap of unknown length
	*	196469	203759:	contig of unknown length
	*	203760	203859:	contig of 7291 bp in length
	*	203860	210741:	gap of unknown length
	*	210742	210841:	contig of 6882 bp in length
	*	210842	216856:	contig of 6015 bp in length
	*	216857	216956:	gap of unknown length
	*	216957	222789:	contig of 5833 bp in length
	*	222790	222889:	gap of unknown length
	*	222890	228284:	contig of 5395 bp in length
	*	228285	228384:	gap of unknown length
	*	228385	229974:	contig of 1590 bp in length
	*	229975	230074:	gap of unknown length
	*	230075	231259:	contig of 1185 bp in length
	*	231260	231359:	gap of unknown length
	*	231360	232369:	contig of 1010 bp in length
	*	232370	232459:	gap of unknown length
	*	232470	233528:	contig of 1059 bp in length.
FEATURES				
source				
			1..233528	
			/organism="Rattus norvegicus"	
			/db_xref="taxon:10116"	
			/chromosome="Mcs1"	
			/clone="CH230-69114"	
BASE COUNT		70468 a	48727 c	46814 g 65514 t 2005 others
ORIGIN				
Query Match			14.2%;	Score 37.6; DB 2; Length 233528;
Best Local Similarity			56.5%;	Pred. No. 1.8;
Matches		70; Conservative	0; Mismatches	54; Indels 0; Gaps 0;
OY	80	tctgccccagagatctcgcttcctcaacctcaaaccttccatcgcgtttaacaca	139	
D	232935	TCTCATTCACACTTTCCTCCCTTTTCCTCCCTCCCTCCCTCCCTTGACACTTACTGATC	232994	
OY	140	aactaccatccgcttcctccctcgatccaatcaactcgcgtccctcttaaacctctcag	199	
D	232995	CCCTTTCCTTTGTCTATTCTCTCTGCCCACTTCCTCCCTCCCTATTATCTCTTATC	233054	
OY	200	gatc	203	
D	233055	GCTC	233058	
RESULT 14				
AF286112/c				
TLOCUS			139005 bp	DNA linear HTG 27-JUN-2001
DEFINITION			Homo sapiens chromosome 8 clone XX-CTB401G11 map 8q, WORKING DRAFT	
ACCESSION			AF286112	SEQUENCE, 12 unordered pieces.
VERSION			AF286112.2	GI:14280189
KEYWORDS			HTG; HTGS_PHASE1; HTGS_DRAFT.	
SOURCE			human.	
ORGANISM			Homo sapiens	
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
			Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
REFERENCE				
AUTHORS			Schilhaber,M.B., Baungart,C., Bleischmidt,K., Dette,M., Jahn,N.,	
			Lehmann,R., Menzel,U., Polley,A., Reichwald,K., Schudy A.,	
			Siddiqui,R., Taudien,S., Wen,G., Siebert,R., Schlegelberger,B.,	
			Rosenthal,A. and Platzer.M.	
			Chromosome 8 genomic sequence	
TITLE			Unpublished	
JOURNAL			2 (bases 1 to 139005)	
REFERENCE			Genome Sequencing Center Jena.	

```

TITLE Direct Submission
JOURNAL Biotechnology, Bielefeldstr. 11, Jena 07745, Germany
COMMENT On Jun 2, 2001 this sequence version replaced gt:1006328.

-----
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj.subm@genome.imb-jena.de
Project Information
-----
Center project name: H201
Center clone name: XX-CTBA01G11
-----
Summary Statistics
Sequencing vector: M13; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 123959 bases at least Q40
Consensus quality: 130354 bases at least Q30
Consensus quality: 134211 bases at least Q20
Quality coverage: 5.19 x in Q20 bases; sum-of-contigs
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 1394: contig of 1394 bp in length
* 1395 1494: gap of unknown length
* 1495 2571: contig of 1077 bp in length
* 2572 2671: gap of unknown length
* 2672 5155: contig of 2484 bp in length
* 5156 5255: gap of unknown length
* 5256 10690: contig of 5435 bp in length
* 10691 10791: gap of unknown length
* 10792 16936: contig of 6146 bp in length
* 16937 17036: gap of unknown length
* 17037 25599: contig of 8563 bp in length
* 25600 25699: gap of unknown length
* 25700 38332: contig of 12623 bp in length
* 38333 38422: gap of unknown length
* 38423 51521: contig of 13099 bp in length
* 51522 51621: gap of unknown length
* 51622 83189: contig of 31566 bp in length
* 83190 83289: gap of unknown length
* 83290 88183: contig of 4894 bp in length
* 88184 88283: gap of unknown length
* 88284 134945: contig of 46662 bp in length
* 134946 135045: gap of unknown length
* 135046 139005: contig of 3960 bp in length.
*
Location/Qualifiers
1..139005
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8g"
/clone="XX-CTBA01G11"
BASE COUNT 41389 a 26216 c 27172 g 43128 t 1100 others
ORIGIN

```


[illegible]

repeat_region	/rpt.family="LIM1" /evidence.not_experimental 5344. .5470 /rpt.family="LIMD1" /evidence.not_experimental complement(.5496. .5613) /rpt.family="FLM.C" /evidence.not_experimental 5991. .6042 /rpt.family="(TA)" /evidence.not_experimental 6042. .6065 /rpt.family="(CA)" /evidence.not_experimental 6120. .7933 /rpt.family="LIPAL5" /evidence.not_experimental 8702. .8773 /rpt.family="AT_rich" /evidence.not_experimental 8979. .9028 /rpt.family="AT_rich" /evidence.not_experimental 9558. .9583 /rpt.family="(TTC)" /evidence.not_experimental 9606. .10112 /rpt.family="AlusX" /evidence.not_experimental complement(.11506. .11612) /rpt.family="MIR" /evidence.not_experimental complement(.13017. .13076) /rpt.family="L2" /evidence.not_experimental complement(.13242. .13542) /rpt.family="Alusg" /evidence.not_experimental 13733. .14012 /rpt.family="AluAb" /evidence.not_experimental 14270. .14583 /rpt.family="Alusg" /evidence.not_experimental 15423. .15677 /rpt.family="AlusX" /evidence.not_experimental 16433. .16486 /rpt.family="Alu" /evidence.not_experimental 16621. .16937 /rpt.family="AlusC" /evidence.not_experimental 17536. .17720 /rpt.family="AluJo" /evidence.not_experimental complement(.18041. .18136) /rpt.family="MIR" /evidence.not_experimental complement(.18538. .18652) /rpt.family="L2" /evidence.not_experimental 19037. .19068 /rpt.family="(A)n" /evidence.not_experimental 19461. .19508 /rpt.family="AT_rich" /evidence.not_experimental 19816. .19836 /rpt.family="AT_rich" /evidence.not_experimental complement(.19891. .20371) /rpt.family="LIME1"
---------------	--


```
repeat_region /evidence=not_experimental
                /rpt_family="AluYb8"
                /evidence=not_experimental
repeat_region complement(20372..20687)
                /rpt_family="AluYb8"
                /evidence=not_experimental
repeat_region complement(20688..20757)
                /rpt_family="L1ME1"
                /evidence=not_experimental
repeat_region 20761..21172
                /rpt_family="MSRB"
                /evidence=not_experimental
repeat_region complement(21248..21941)
                /rpt_family="L1M4"
                /evidence=not_experimental
repeat_region complement(21947..22328)
                /rpt_family="L1M4"
                /evidence=not_experimental
repeat_region 22329..22629
                /rpt_family="AluSx"
                /evidence=not_experimental
repeat_region complement(22630..23848)
                /rpt_family="L1M4"
                /evidence=not_experimental
repeat_region complement(23869..23980)
                /rpt_family="AluJc/FRAM"
                /evidence=not_experimental
repeat_region 24028..24058
                /rpt_family="AT_rich"
                /evidence=not_experimental
repeat_region complement(24059..24231)
                /rpt_family="L1M4"
                /evidence=not_experimental
repeat_region complement(24232..24544)
                /rpt_family="AluSg"
                /evidence=not_experimental
repeat_region complement(24545..24643)
                /rpt_family="L1M4"
                /evidence=not_experimental
repeat_region complement(25158..25288)
                /rpt_family="MIR"
                /evidence=not_experimental
repeat_region 26569..26660
                /rpt_family="MIR"
                /evidence=not_experimental
repeat_region 26955..27007
                /rpt_family="(TA)n"
                /evidence=not_experimental
repeat_region 27520..27551
                /rpt_family="(T)n"
                /evidence=not_experimental
repeat_region complement(27852..28148)
                /rpt_family="L1PA15"
                /evidence=not_experimental
repeat_region complement(28205..28311)
                /rpt_family="MIR"
                /evidence=not_experimental
repeat_region complement(28488..28760)
                /rpt_family="L2"
                /evidence=not_experimental
repeat_region complement(29003..29289)
                /rpt_family="AluSc"
                /evidence=not_experimental
repeat_region complement(29457..29758)
                /rpt_family="AluSc"
                /evidence=not_experimental
repeat_region complement(30384..30452)
                /rpt_family="L2"
                /evidence=not_experimental
repeat_region 30869..30922
                /rpt_family="AT_rich"
                /evidence=not_experimental
repeat_region 31792..31814
                /rpt_family="AT_rich"
                /evidence=not_experimental
```

```
Query Match      14.2%; Score 37.4; DB 9; Length 165901;
Best Local Similarity 52.2%; Pred. No. 2.2;
Matches 83; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 32 ttccaagctctcaacgccaatccacgtgcttcccttaccagaagcgctcgcccgaga 91
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 44233 TCCAAGGTTATCATCAGACAGATTACATCAGTATCATCAGACAGACTTCCATCGGGT 44292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92 ttctcgcttcccaacctccaaccttctccatcgcttcttaatacaaaactaacctatc 151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 44293 TTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 44352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 152 cgcctccctcgatccatccatccatcgctccctccctttaaac 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 44353 CTCCTCCCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAAC 44391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: July 3, 2002, 05:17:46
Job time: 49246 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 07:58:10 ; Search time 796.36 Seconds

(without alignments)
569.171 Million cell updates/sec

Title: US-09-300-482-569

Perfect score: 264
Sequence: 1 ctgactgcagcaattccca.....gacatgcttcattgcttta 264

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_037802:*
1: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	34.8	13.2	7329	22	AA546674	Tumour suppressor
c 2	34.8	13.2	107820	22	AA016230	Human ATP-binding
c 3	34	12.9	27082	22	AA070447	Human immune/haema
c 4	33.2	12.6	5647	24	ABL33567	Human immune syste
c 5	33.2	12.6	5647	24	AA561321	Human gene regulat
c 6	33.2	12.6	6531	24	ABL32640	Human immune syste
c 7	33	12.3	9007	22	AA546615	Tumour suppressor
c 8	32.6	12.3	5942	24	AA561134	Human gene regulat
c 9	32.4	12.3	857	20	AA57401	Rat U3 gene trap d

c 10	32.4	12.3	14032	24	ABL33453	Human immune syste
c 11	32	12.1	7038	22	AA545495	Chemically pretrea
c 12	32	12.1	7038	24	ABL34096	Human immune syste
c 13	31.8	12.0	1684	24	ABL01534	Murine apoptosis r
c 14	31.8	12.0	5565	24	ABL32262	Human immune syste
c 15	31.8	12.0	8649	22	AA546330	Tumour suppressor
c 16	31.6	12.0	460	22	AA539053	Novel human diagn
c 17	31.6	12.0	16748	22	AA651193	Human immune/haema
c 18	31.4	11.9	1612	19	AAV44595	Human respiration
c 19	31.4	11.9	4002	21	AA488575	Human Son of seven
c 20	31.4	11.9	4002	21	AA60603	Nucleotide sequenc
c 21	31.4	11.9	13605	22	AA137081	Human musculoskele
c 22	31.2	11.8	5798	24	ABL32168	Human immune syste
c 23	31.2	11.8	53526	19	AA794101	Human PKD1 gene.
c 24	31.2	11.8	53577	17	AA718551	Human polycystic k
c 25	31.2	11.8	53577	19	AA794108	Human PKD1 locus b
c 26	30.8	11.7	675	24	AA562171	Porcine muscular s
c 27	30.8	11.7	10286	22	AA545308	Chemically pretrea
c 28	30.8	11.7	36901	20	AA23892	Murine LDBO genom
c 29	30.8	11.7	38886	20	AA23897	Murine LDBO homolo
c 30	30.6	11.6	1543	21	AA707595	Fusarium venenatum
c 31	30.6	11.6	3837	23	AA74708	DNA encoding novel
c 32	30.6	11.6	8333	22	AA545407	Chemically pretrea
c 33	30.6	11.6	8333	24	ABL33503	Human immune syste
c 34	30.6	11.6	31562	23	ABL03868	Drosophila melanog
c 35	30.6	11.6	32169	22	ABA14358	Human nervous syst
c 36	30.4	11.5	399	22	AA586561	Novel human diagn
c 37	30.2	11.4	541	23	AA586561	DNA encoding novel
c 38	30.2	11.4	862	22	AA194410	Human neuroblastom
c 39	30.2	11.4	5270	24	ABL333039	Human immune syste
c 40	30.2	11.4	6656	24	ABL33400	Human immune syste
c 41	30.2	11.4	8212	24	ABL99884	Mouse ischaemic co
c 42	30.2	11.4	15714	22	AA536855	Human cardiovascular
c 43	30	11.4	6310	24	AA561269	Human gene regulat
c 44	30	11.4	6807	22	ABA19437	Human nervous syst
c 45	30	11.4	7011	24	ABL32547	Human immune syste

ALIGNMENTS

RESULT 1
ID AA546674 standard; DNA: 7329 BP.
XX
AC AA546674;
XX
DT 18-DEC-2001 (first entry)
XX
DE Tumour suppressor gene derived chemically modified sequence #396.
XX
KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; Cdc dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
XX
OS Homo sapiens.
XX
PN WC0200168912-A2.
XX
PD 20-SEP-2001.
XX
PE 15-MAR-2001; 2001WO-EP02955.
XX
PR 15-MAR-2000; 2000DE-1013847.
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPig-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX

DR WPI: 2001-602752/68.

XX Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer

PS Claim 1; SEQ ID No 396; 27pp; English.

XX
CC The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (ss) and sequences complementary to (ss). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes. Sequences with even numbered Seq ID numbers are the
CC complementary sequence of the corresponding odd numbered sequence (e.g.
CC ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
CC is missing).

CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 7329 BP; 1917 A; 141 C; 1597 G; 3670 T; 4 other;

Query Match 13.2%; Score 34.8; DB 22; Length 7329;
Best Local Similarity 50.6%; Pred. No. 0.45;

Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 32 ttcaagctctaagccaatcgaactgctctcattcagaagcgtccgcccgaga 91
Db 696 TTTCGAAACCTCAAAATTAATTAATTTCTCTTACCAAAAACCTTCCCATAAA 637
Qy 92 ttctgcctctcaacctcaacctcttcctcatgctttatcaacaacctatc 151
Db 636 TACTTACTCAACCTTAAAAATCTCAACACCTCAAAACTTAATAATTCACAAAA 577
Qy 152 cgtctccctccgatactcaactgcgtccctcttaaacctcctta 197
Db 576 CTCTTCAACACATCCACCAATATCTTCTTAATAAAGTCCCTCTA 531

RESULT 2

AAD16230
ID AAD16230 standard; DNA: 107820 BP.

XX
AC AAD16230;

XX
DT 19-NOV-2001 (first entry)

XX
DE Human ATP-binding cassette transporter ABCG6 (MRP6) complementary gene.

KW Human: prenatal diagnosis; dermal lesion; cardiovascular disease; MRP6;
KW Multidrug Resistance-associated protein 6; macular degeneration; ABCG6;
KW ATP-binding cassette transporter; arterial insufficiency; chromosome 16;
KW Pseudoxanthoma elasticum; PXE; heritable disorder; retinal haemorrhage;
KW ss.

XX
OS Homo sapiens.

XX	Key	Location/Qualifiers
FX	CDS	complement (102783..29773)
FT		/*tag= a
FT		/product= "Human ABCG6 (MRP6) protein"
FT		complement (107820..102784)
FT	intron	/*tag= b
FT		/number= 1
FT	exon	complement (102783..102748)
FT		/*tag= c
FT	intron	/number= 1
FT		complement (102747..101181)
FT		/*tag= d
FT	exon	/number= 2
FT		complement (101180..100998)
FT		/*tag= e
FT	intron	/number= 2
FT		complement (100997..99297)
FT		/*tag= f
FT	exon	/number= 3
FT		complement (99296..99171)
FT		/*tag= g
FT	intron	/number= 3
FT		complement (99170..99032)
FT		/*tag= h
FT	exon	/number= 4
FT		complement (99031..98903)
FT		/*tag= i
FT	intron	/number= 4
FT		complement (98902..93799)
FT		/*tag= j
FT	exon	/number= 5
FT		complement (93798..93676)
FT		/*tag= k
FT	intron	/number= 5
FT		complement (93675..91595)
FT		/*tag= l
FT	exon	/number= 6
FT		complement (91594..91533)
FT		/*tag= m
FT	intron	/number= 6
FT		complement (91532..88208)
FT		/*tag= n
FT	exon	/number= 7
FT		complement (88207..88076)
FT		/*tag= o
FT	intron	/number= 7
FT		complement (88075..82955)
FT		/*tag= p
FT	exon	/number= 8
FT		complement (82954..82757)
FT		/*tag= q
FT	intron	/number= 8
FT		complement (82756..81525)
FT		/*tag= r
FT	exon	/number= 9
FT		complement (81524..81347)
FT		/*tag= s
FT	intron	/number= 9
FT		complement (81346..77529)
FT		/*tag= t
FT	exon	/number= 10
FT		complement (77528..77367)
FT		/*tag= u
FT	intron	/number= 10
FT		complement (77366..72269)
FT		/*tag= v
FT	exon	/number= 11
FT		complement (72268..72176)
FT		/*tag= w
FT	intron	/number= 11
FT		complement (72175..69719)
FT		/*tag= x

FT	exon	/number= 12 /complement (69718..69515) /*tag= y /number= 12 complement (69514..68326) /*tag= z
FT	exon	/number= 13 /complement (68325..68182) /*tag= aa /number= 13 complement (68181..66563) /*tag= ab
FT	exon	/number= 14 /complement (66562..66475) /*tag= ac /number= 14 complement (66474..64386) /*tag= ad
FT	exon	/number= 15 /complement (64385..64310) /*tag= ae /number= 15 complement (64309..62283) /*tag= af
FT	exon	/number= 16 /complement (62282..62156) /*tag= ag /number= 16 complement (62155..61941) /*tag= ah
FT	exon	/number= 17 /complement (61940..61764) /*tag= ai /number= 17 complement (61763..58325) /*tag= aj
FT	exon	/number= 18 /complement (58324..58157) /*tag= ak /number= 18 complement (58156..56986) /*tag= al
FT	exon	/number= 19 /complement (56985..56811) /*tag= am /number= 19 complement (56810..55346) /*tag= an
FT	exon	/number= 20 /complement (55345..55270) /*tag= ao /number= 20 complement (55269..52758) /*tag= ap
FT	exon	/number= 21 /complement (52757..52637) /*tag= aq /number= 21 complement (52636..49589) /*tag= ar
FT	exon	/number= 22 /complement (49588..49381) /*tag= as /number= 22 complement (49380..45579) /*tag= at
FT	exon	/number= 23 /complement (45578..45268) /*tag= au /number= 23 complement (45267..42838) /*tag= av
FT	exon	/number= 24

FT	exon	complement (42837..42638)
FT		/*tag= aw
FT		/number= 24
FT	intron	complement (42637..41210)
FT		/*tag= ax
FT		/number= 25
FT	exon	complement (41209..41083)
FT		/*tag= ay
FT		/number= 25
FT	intron	complement (41082..39227)
FT		/*tag= az
FT		/number= 26
FT	exon	complement (39226..39125)
FT		/*tag= ba
FT		/number= 26
FT	intron	complement (39124..37454)
FT		/*tag= bb
FT		/number= 27
FT	exon	complement (37453..37307)
FT		/*tag= bc
FT		/number= 27
FT	intron	complement (37306..34675)
FT		/*tag= bd
FT		/number= 28
FT	exon	complement (34674..34516)
FT		/*tag= be
FT		/number= 28
FT	intron	complement (34515..34438)
FT		/*tag= bf
FT		/number= 29
FT	exon	complement (34437..34271)
FT		/*tag= bg
FT		/number= 29
FT	intron	complement (34270..30413)
FT		/*tag= bh
FT		/number= 30
FT	exon	complement (30412..30218)
FT		/*tag= bi
FT		/number= 30

Query Match	13.28;	Score 34.8;	DB 22;	Length 107820;
Best Local Similarity	50.0%;	Pred. No. 1.3;		
Matches 87; Conservative	0;	Mismatches 87;	Indels 0;	Gaps 0;

[illegible]

RESULT

ID AAK70447 standard; DNA; 27082 BP.

AC AAK70447

DT 06-NOV-2001 (first entry)

DE	Human immune/haematopoietic antigen genomic sequence	SEQ ID NO:25259.
....		

Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

KW cytostatic; gene therapy; vaccine; metastasis; ds

OS Homo sapiens.

PN WO200157182-A2.

XX 09-AUG-2001.
PD 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0203515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220563.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-022547.
PR 14-AUG-2000; 2000US-022557.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229345.
PR 01-SEP-2000; 2000US-0229509.
PR 03-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
XX

PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237040.
PR 02-OCT-2000; 2000US-0237043.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241281.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI

DR	WT: 2001-483426/52.	
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PR	useful for preventing, diagnosing and/or treating cancers and	
PT	metastasis -	
XX		
PS	Disclosure: SEQ ID NO 25259; 3071bp + Sequence Listing; English.	
XX		
CC	AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)	
CC	amino acid sequences given in AAK82170 to AAK91921. (I) have cytosolic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins, and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patients own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting	
CC	the nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/hematopoietic-related diseases, especially	
CC	cancers and cancer metastases of hematopoietic-derived cells; AAK64703	
CC	to AAK87694 represent human immune/hematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAK82169	
CC	represent sequences used in the exemplification of the present invention.	
XX		
SQ	Sequence 27082 BP; 5963 A; 7786 C; 7711 G; 5622 T; 0 other;	
	Query Match	12.9%; Score 34; DB 22; Length 27082;
	Best Local Similarity	61.1%; Pred. No. 1.4;
	Matches	55; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
OY	113 acctcttcacatcgcttttaacacaaactaacctacgcgttcctccgatcatcca 172	
Db	1676 agcttcgcgcacatcttccaaagcccaactcaactcactcctccgcagcgtcccca 1735	
OY	173 ctgcctccctcttaaacctccttacggat 202	
Db	1736 catgcccttcattcttccttccttgat 1765	
	RESULT	
	4	
ABL33567/C		
ID	ABL33567 standard; DNA; 5647 BP.	
XX		
AC	ABL33567;	
XX		
DT	26-MAR-2002 (first entry)	
XX		
DE	Human immune system associated gene SEQ ID NO: 1540.	
XX		
XX	Human: immune system disease; cytosine methylation; antiasthmatic;	
KW	antiartherosclerotic; anlianaemic; cytosolic; noctropic;	
KW	neutropoietic; anti-HIV; anticonvulsant; ophthalmological;	
KW	antiinflammatory; antiarthritic; antidiabetic; antipsoriatic;	
KW	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;	
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;	
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;	
KW	gene; ds.	
KW		
OS	Homo sapiens.	
XX		
PN	WO200200928-A2.	
XX		
PD	03-JAN-2002.	
XX		
PF	02-JUL-2001; 2001WO-EP07537.	
XX		
PR	30-JUN-2000; 2000DE-1032529.	
PR	01-SEP-2000; 2000DE-1043826.	
XX		
PA	(EPIG-) EPIGENOMICS AG.	

XX	Olek A, Piepenbrock C, Berlin K;
PI	WPI: 2002-130909/17.
DR	
XX	
PT	Nucleic acid comprising fragment of chemically modified gene, useful
PT	for diagnosis and treatment of diseases associated with abnormal
PT	cytosine methylation
XX	
PS	Claim 1; SEQ ID NO 1540; 32pp + Sequence Listing; German.
XX	
CC	The present invention provides a number of human immune system associated
CC	genes which are modified by the methylation of cytosines. The sequences
CC	can be used in the diagnosis and treatment of immune system disorders,
CC	including eye diseases such as retinopathy, neovascular glaucoma and
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC	leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC	diseases. The present sequence is a gene of the invention.
XX	
SO	Sequence 5647 BP: 2105 A; 52 C; 1149 G; 2341 T; 0 other;
Query Match	12.6%; Score 33.2; DB 24; Length 5647;
Best Local Similarity	51.4%; Pred. No. 1.4;
Matches 109; Conservative	0; Mismatches 93; Indels 10; Gaps 1
QY	39 ctctcaagcccaatcactgtcttccctatcagaagcgcttcgccccgagattctgc 98
DB	943 CTATCAATTCATCAATTAACAACACTATTTCTCAAAAAAACATTAACCAATTTCTCTC 884
QY	99 ttccctacctcaaacctctccactgcgttttaactacaacaaactaatccgcttc 158
DB	883 CTCCTTCCCCCTTAACCCCTCTCCATTTATCTAAGTAACGTAACCTTTACGATTTA 824
QY	159 ctccgatacca-----tctactgcctccctttaaacctccttaagatacaatg 208
DB	823 ATATTATACAACTTCCACATTTCCCTCCGCTCCACCAACCCCTCCCAAAATT 764
QY	209 ctcccaacaagatggagatggaatgcctgcaa 240
DB	763 CAAAAAACTCCCGTTACCGCAATATACCA 732
RESULT 5	
AAS61321/C	
ID	AAS61321 standard; DNA; 5647 BP.
XX	
AC	AAS61321;
XX	
DT	29-JAN-2002 (first entry)
XX	
DE	Human gene regulation-associated gene oligonucleotide #276.
XX	
KW	Human; Gene regulation-associated gene; severe combined immunodeficiency;
KW	cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
KW	asthma; HIR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
KW	renal disease; Preeclampsia; cardiac allograft vascular disease;
KW	colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
KW	immunostimulant; cardiac; antiinflammatory; coagulant; antiasthmatic;
KW	neurotropic; gynecological; anti-tumour; immunosuppressive; cystostatic.
XX	
OS	Homo sapiens.
XX	
PN	WO200177375-A2.
XX	
PD	18-OCT-2001.
XX	
PF	06-APR-2001; 2001WO-EP03968.
XX	
PR	06-APR-2000; 2000DE-1019058.
PR	07-APR-2000; 2000DE-1019173.
PR	30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.
 XX (EPiG-) EPIGENOMICS AG.
 XX Olek A, Piepenbrock C, Berlin K;
 XX WPI: 2002-017470/02.
 DR
 XX
 PT New nucleic acid sequences from chemically modified genes associated
 PT with gene regulation, useful for analysing cytosine methylations for
 PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
 PT disease
 PS Disclosure: SEQ ID No 282; 26pp; English.
 XX
 XX The invention relates to 224 nucleic acid sequences comprising at least
 CC 18 bases of a chemically pretreated gene associated with gene regulation
 CC selected from 43 known genes (or complementary sequences). The
 CC chemical pretreatment converts cytosine bases unmethylated at the
 CC 5-position to uracil or another base with hybridisation behaviour
 CC dissimilar to cytosine, to enable analysis of cytosine methylations.
 CC The DNA sequences, oligomers (or sets/arrays) and method are
 CC useful in the diagnosis of diseases (or predisposition to diseases)
 CC associated with gene regulation and in therapy of such diseases, by
 CC enabling analysis of the cytosine methylation patterns of such genes,
 CC kits are provided. They are especially useful in diagnosis
 CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
 CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
 CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
 CC preclampsia, graft versus-host disease. The present sequence is a
 CC sequence included in the sequence data for this specification and is
 CC associated with the human gene regulation-associated genes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 5647 BP; 2105 A; 52 C; 1149 G; 2341 T; 0 other;

Query Match 12.6%; Score 33.2; DB 24; Length 5647;
 Best Local Similarity 51.4%; Pred. No. 1.4;
 Matches 109; Conservative 0; Mismatches 93; Indels 10; Gaps 1;

QY 39 cttcaagcgcaatccactgcttctcattcagaagcgcttcgcccgaagatctcgc 98
 DB 943 CTATCAATTCAATTAACAACATAATCTCAAAAAACAATTAACCAATTTCTCTC 884
 QY 99 ttctcacttcaaaccttctccatcgctttaaatacacaactaactacgcgtcc 158
 DB 883 CTCTTCCCTCTAACCCCTCTCTCAATTATATCTAATCAACGTAACCTTTACGATTGA 824
 QY 159 ctccgagatca-----ttcactcgctccctcttaaacctcttaagatgaatg 208
 DB 823 ATATTATACAACTTCACATTTCCCTCGCTCCACACCAACCCCTCCCAACCAATTT 764
 QY 209 ctcccaaacagatgggaagatcctgcaa 240
 DB 763 CAAAAAAACTCCCTTTACCGCAATACACCAA 732

RESULT 6
 ABL32640/C
 ID ABL32640 standard; DNA: 6531 BP.
 XX
 XX ABL32640;
 AC
 XX
 XX 26-MAR-2002 (first entry)
 DT
 XX
 XX Human immune system associated gene SEQ ID NO: 613.
 DE
 XX
 XX Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiasthenosclerotic; antianaemic; cytosinatic; nootropic;
 KW

KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antiproliferic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 OS Homo sapiens.
 XX
 XX WO200200928-A2.
 PN
 XX
 PD 03-JAN-2002.
 XX
 XX 02-JUL-2001; 2001WO-EP07537.
 PF
 XX 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 XX (EPiG-) EPIGENOMICS AG.
 PA
 PI Olek A, Piepenbrock C, Berlin K;
 DR WPI: 2002-130909/17.
 XX
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation
 PS Claim 1: SEQ ID NO 613; 32pp + Sequence listing; German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 CC
 SQ Sequence 6531 BP; 1840 A; 131 C; 1648 G; 2912 T; 0 other;

Query Match 12.6%; Score 33.2; DB 24; Length 6531;
 Best Local Similarity 57.8%; Pred. No. 1.5;
 Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 98 cttccacttcaaaccttctccatcgctttaaatacacaactaactatcgcttc 157
 DB 4980 CTTACTCGAATCGAAATCCCTCTCTCCGCAACGAAACCTCAATTAACCAATCCGCTC 4921
 QY 158 ctccgagatccatccactcgctccctcttaaacctccttaag 199
 DB 4920 GCGCGACCAATTCCTACTACTAGCGAAATACACCAACG 4879

RESULT 7
 AAS46615/C
 ID AAS46615 standard; DNA: 9007 BP.
 XX
 XX AAS46615;
 AC
 XX
 XX 18-DEC-2001 (first entry)
 DT
 XX
 XX Tumour suppressor gene derived chemically modified sequence #337.
 DE
 XX
 XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
 KW cancer; tumour; Cpg dinucleotide; single-nucleotide polymorphism; SNP;
 KW cytosine methylation; ds.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200168912-A2.
 PN
 XX

PD 20-SEP-2001.
XX
PF 15-MAR-2001; 2001WO-EP02955.
XX
PR 15-MAR-2000; 2000DE-1013847.
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-602752/68.
XX
PT Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
cancer
XX
PS Claim 1; SEQ ID No 337; 27pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (ss) and sequences complementary to (ss). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 9007 BP; 1750 A; 464 C; 2703 G; 4090 T; 0 other;
Query Match 12.5%; Score 33; DB 22; Length 9007;
Best Local Similarity 58.8%; Pred. No. 1.9;
Matches 57; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 94 ctgcctctactcttaaacctcttcacgtcttcaacacaaactaacctatccg 153
DB 576 CCGGACCCCTCTCTCCACACCTTCCTTCGACTCAACAAACAAACCTCTCT 517
QY 154 ctccctccgcagcatcactcgcctccctcttaaac 190
DB 516 TTCTAAATCCTCTTTTCACCTACTACCAACACACAC 480
RESULT 8
AAS61134/c
ID AAS61134 standard; DNA: 5942 BP.
XX
AC AAS61134;
XX
DT 29-JAN-2002 (first entry)
XX
DE Human gene regulation-associated gene oligonucleotide #89.

KW Human; Gene regulation-associated gene; severe combined immunodeficiency;
KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
KW renal disease; Preeclampsia; cardiac allograft vascular disease;
KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
KW immunostimulant; cardiac; antiinflammatory; coagulant; antithrombotic;
KW nephrologic; gynecological; anti-tumour; immunosuppressive; cytostatic.
XX
OS Homo sapiens.
XX
PN WO200177375-A2.
XX
PD 18-Oct-2001.
XX
PF 06-APR-2001; 2001WO-EP03968.
XX
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-017470/02.
XX
PT New nucleic acid sequences from chemically modified genes associated
PT with gene regulation, useful for analysing cytosine methylations for
PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
PT disease
XX
PS Claim 1; SEQ ID No 91; 26pp; English.
XX
CC The invention relates to 224 nucleic acid sequences comprising at least
CC 18 bases of a chemically pretreated gene associated with gene regulation
CC selected from 43 known genes (or complementary sequences). The
CC chemical pretreatment converts cytosine bases unmethylated at the
CC 5-position to uracil or another base with hybridisation behaviour
CC dissimilar to cytosine, to enable analysis of cytosine methylations.
CC The DNA sequences, oligomers (or sets/arrays) and method are
CC useful in the diagnosis of diseases (or predisposition to diseases)
CC associated with gene regulation and in therapy of such diseases, by
CC enabling analysis of the cytosine methylation patterns of such genes,
CC kits are provided. They are especially useful in diagnosis
CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
CC preeclampsia, graft versus-host disease. The present sequence is a
CC sequence included in the sequence data for this specification and is
CC associated with the human gene regulation-associated genes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5942 BP; 1401 A; 180 C; 1593 G; 2768 T; 0 other;
Query Match 12.3%; Score 32.6; DB 24; Length 5942;
Best Local Similarity 58.0%; Pred. No. 2.2;
Matches 76; Conservative 0; Mismatches 54; Indels 1; Gaps 1;
QY 86 ccgagatctcgcctccacacttcaaacctcttcacgtcttcaacacaaactaa 145
DB 4876 CCGGACACTCTCTCAGCTTCCTTAATCCCGAGCGACCGCTACTACCAAAAAACATA 4817
QY 146 cctatcgcctccctccgcatcactcgcctccctcttaaacctccctctagatcaa 205
DB 4816 CCTATCCTCGACCAATCCGTAATTC-CTTATTAAGCTTTAACCCTTAATAGCAATCA 4758
QY 206 atgtctcccaaa 216

[illegible]

XX	26-MAR-2002	(first entry)	
DT			
XX			
XX	Human immune system associated gene	SEQ ID NO: 2069.	
DE			
XX	Human; immune system disease; cytosine methylation; antiasthmatic;		
KW	antiartherosclerotic; antihaemic; cytosatic; nootropic;		
KW	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;		
KW	antihemmatic; antiarthritic; antidabetic; antipsoriatic;		
KW	antihemorrhagic; cancer; eye disease; arteriosclerosis; anaemia;		
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;		
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;		
XX	gene; ds.		
XX			
OS	Homo sapiens.		
XX			
XX	MO200200928-A2.		
PD			
XX	03-JAN-2002.		
XX			
XX	02-JUL-2001; 2001WO-EP07537.		
XX			
XX	30-JUN-2000; 2000DE-1032529.		
PR	01-SEP-2000; 2000DE-1043826.		
XX			
PA	(EPIC-) EPIGENOMICS AG.		
XX			
PI	Olek A, Piepenbrock C, Berlin K;		
XX			
DR	WPI: 2002-130909/17.		
XX			
PT	Nucleic acid comprising a fragment of chemically modified gene, useful		
PT	for diagnosis and treatment of diseases associated with abnormal		
PT	cytosine methylation		
XX			
PS	Claim 1; SEQ ID NO 2069; 32pp + Sequence Listing; German.		
XX			
CC	The present invention provides a number of human immune system associated		
CC	genes which are modified by the methylation of cytosines. The sequences		
CC	can be used in the diagnosis and treatment of immune system disorders,		
CC	including eye diseases such as retinopathy, neovascular glaucoma and		
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid		
CC	leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,		
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel		
CC	diseases. The present sequence is a gene of the invention.		
XX			
SQ	Sequence 7038 BP; 1390 A; 375 C; 2162 G; 3111 T; 0 other;		
XX			
XX	Query Match	12.1%; Score 32; DB 24; Length 7038;	
XX	Best Local Similarity	50.7%; Pred. No. 3.7; Indels	0; Gaps 0;
XX	Matches	77; Conservative	0; Mismatches 75;
XX			
QY	69	tcagaagcgcttcgcccgcgagatcctcgctccctcaactcaaacctctccatcgct	128
DB	2981	TAAAAACACGACTCTTAATAAAACCCCTCCATCCACCTACTAATAATTCTTCCCT	2922
QY	129	tttaatacaaaaactaacctatcgctcctccctcgatcattcaactcgctccctcttaa	188
DB	2921	ATCGGCACACCAATTAATCTTAATCTTACCTCAATCTCCCTACAACTACCACTTAA	2862
QY	189	acctcttagcgatcaaatgctcccaaacaga	220
DB	2861	AAAACCTAACCCACGAAAACCTCTTAACACA	2830
XX			
XX	RESULT 13		
XX	ABL01534		
XX	ABL01534 standard; DNA; 1684 BP.		
XX	ABL01534;		
XX			
XX	15-MAR-2002 (first entry)		
DT			

XX DE Murine apoptosis related DNA sequence #199.
XX

OS Apoptosis; mouse; cancer; autoimmune disease; viral infection;
PN Alzheimer's disease; Parkinson's disease; Huntington's disease;
XX reperfusion injury; stroke; liver damage; dilatory cardiomyopathy;
KW transgenic animal; hepatotropic; antialcoholism; cytoskeletal;
KW immunosuppressive; virucide; neuroprotective; vasotropic;
XX antiparkinsonian; cerebroprotective; ds.
OS Mus sp.

DE10126344-A1.

24-JAN-2002.

30-MAY-2001; 2001DE-1026344.

14-JUL-2000; 2000DE-1034303.

(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PI Grimm S, Schoenfeld N, Braziulis E, Cramer U, Gewies A, Voss F,
PI Mund T, Albayrak T, Gille H, Klein M;

WPI; 2002-115563/16.

PT New apoptosis-associated nucleic acid sequences and polypeptides,
PT useful for diagnosis, treatment and prevention of e.g. tumors and
XX neurodegeneration -

Claim 1; Page 214-215; 227pp; German.

CC The present invention relates to nucleic acids from the mouse, where the
CC nucleic acid is associated with apoptosis. The sequences can be used in
CC the diagnosis, treatment and prevention of diseases associated with
CC excessive or inadequate apoptosis, including tumours, autoimmune
CC diseases, viral infections, degenerative diseases (Alzheimer's,
CC Parkinson's and Huntington's diseases), reperfusion injury, stroke and
CC alcohol-induced injury to the liver, for identifying agents for treating
CC these diseases, and to prepare transgenic animals in which expression of
CC an apoptosis related sequence is altered. These are useful for genetic
CC and/or pharmacological investigations of apoptosis and related diseases,
CC including dilatory cardiomyopathy. The present sequence is one of the
CC apoptosis related sequences of the invention.

Sequence 1684 BP; 147 A; 669 C; 52 G; 459 T; 357 other;

Query Match 12.0%; Score 31.8; DB 24; Length 1684;
Best Local Similarity 43.7%; Pred. No. 2.5;

Matches 69; Conservative 0; Mismatches 87; Indels 2; Gaps 1;

QY 31 ttccaagctctcaagcgaatccactgctctctatcagaagcgttcgccccgag 90
DB 787 ttccctctcnaacnmcnctnmcnctnmcnctnmcnctnmcnctnmcnctnmcn 846
QY 91 attcgcgttctcaacttcaaaccttctccatgctt--ttaatcaaaaactaacct 148
DB 847 tctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 148
QY 149 atcgcgtctccctcgcattcattcattcattcattcattcattcattcattcatt 186
DB 907 ncccnctnctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 944

RESULT 14

ABL32262/c

ABL32262 standard; DNA; 5565 BP.

26-MAR-2002 (first entry)

XX DE Human immune system associated gene SEQ ID NO: 225.
XX

OS Human; immune system disease; cytosine methylation; antiaesthetic;
PN antiaesthetic; antiaesthetic; antiaesthetic; antiaesthetic;
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antineumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antineumatic; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
XX gene; ds.

OS Homo sapiens.

WO200200928-A2.

03-JAN-2002.

02-JUL-2001; 2001WO-EP07537.

30-JUN-2000; 2000DE-1032529.

01-SEP-2000; 2000DE-1043826.

(EPIG-) EPIGENOMICS AG.

Olek A, Piepenbrock C, Berlin K;

WPI; 2002-130909/17.

PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
XX cytosine methylation -

Claim 1; SEQ ID NO 235; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.

Sequence 5565 BP; 1405 A; 76 C; 1326 G; 2758 T; 0 other;

Query Match 12.0%; Score 31.8; DB 24; Length 5565;
Best Local Similarity 52.7%; Pred. No. 4;
Matches 69; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 67 tatcagaagcgttcgccccgagattcgtctctcaactcaaacctctccatcg 126
DB 403 TATAAAAATTCCTACTCCACTATATCCCAACCTTCTCATATATCCCTTTCACCAA 344
QY 127 cttaatacaaaaactaacatcgcgttccctcgcattcattcattcattcattcatt 186
DB 343 CCTTAACTCATATCAATTAACATATTTCTTACCACATCTCCAAAAACTTAACCTTCCCA 284
QY 187 aaactcctta 197
DB 283 AAACCTCACTA 273

RESULT 15

AAS46330/c

AAS46330 standard; DNA; 8649 BP.

18-DEC-2001 (first entry)

Tumour suppressor gene derived chemically modified sequence #52.

XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP,
XX cytosine methylation; ds.
XX
XX Homo sapiens.
XX
XX WO200168912-A2.
XX
XX 20-SEP-2001.
XX
XX 15-MAR-2001; 2001WO-EP02955.
XX
XX 15-MAR-2000; 2000DE-1013847.
XX 06-APR-2000; 2000DE-1019058.
XX 07-APR-2000; 2000DE-1019173.
XX 30-JUN-2000; 2000DE-1032529.
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPig-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI: 2001-602752/68.
XX
XX Fragments of chemically modified genes associated with tumour suppressor
XX genes and oncogenes, useful in designing primers and probes for
XX PT analysing diseases associated with cytosine methylation state e.g.
XX PT cancer -
XX
XX Claim 1; SEQ ID No 52; 27pp; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of 18
XX bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
XX bisulphite, of genes associated with tumour suppression and
XX CC oncogenes having a sequence taken from 536 (actually 533 since
XX CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
XX CC (Ss) and sequences complementary to (Ss). The nucleic acid may be a
XX CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
XX CC form part of a set of probes for detecting the cytosine methylation state
XX CC of/or single nucleotide polymorphisms and also to be used in an
XX CC array for analysing diseases associated with CpG dinucleotides e.g.
XX CC cancers and tumours. The probes can also be used in a method for
XX CC ascertaining genetic and/or epigenetic parameters for the diagnosis
XX CC and/or therapy of existing diseases or the predisposition to specific
XX CC diseases, by analysing cytosine methylations. The parameters may be
XX CC compared to another set of genetic and/or epigenetic parameters, the
XX CC differences serving as basis for diagnosis and/or prognosis events which
XX CC are disadvantageous to patients. The present sequence is one of the
XX CC 533 genetic sequences derived from tumour suppressor genes and
XX CC oncogenes. Sequences with even numbered Seq ID numbers are the
XX CC complementary sequence of the corresponding odd numbered sequence (e.g.
XX CC ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
XX CC is missing).
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 8649 BP; 1859 A; 263 C; 2358 G; 4169 T; 0 other;
XX

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 08:11:38 ; Search time 185.98 Seconds
(without alignments)
348.679 Million cell updates/sec

Title: US-09-300-482-569

Perfect score: 264
1 ctgactgtcagcaatcccca.....gacatgtcattcatcttta 264

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51.2	19.4	7218	1 US-08-232-463-14	Sequence 14, Appl
2	31.4	11.9	4002	3 US-09-356-952-9	Sequence 9, Appl
3	31.2	11.8	53526	3 US-08-658-136-2	Sequence 2, Appl
4	31.2	11.8	53577	3 US-08-658-136-1	Sequence 1, Appl
5	30.6	11.6	289	4 US-09-007-005-17	Sequence 17, Appl
6	30.6	11.6	289	4 US-09-244-796-17	Sequence 17, Appl
7	29.2	11.1	12793	4 US-09-004-838-124	Sequence 124, App
8	29	11.0	5751	4 US-09-417-455-7	Sequence 7, Appl
9	29	11.0	5751	4 US-09-348-942-7	Sequence 7, Appl
10	28.8	10.9	2575	1 US-08-135-511-33	Sequence 33, Appl
11	28.8	10.9	2575	1 US-08-483-852-10	Sequence 10, Appl
12	28.8	10.9	2575	1 US-08-361-458-5	Sequence 5, Appl
13	28.8	10.9	2575	1 US-08-477-953-10	Sequence 10, Appl
14	28.8	10.9	2575	1 US-08-187-453-33	Sequence 33, Appl
15	28.8	10.9	2575	2 US-08-477-952-10	Sequence 10, Appl
16	28.8	10.9	36741	4 US-09-301-665-3	Sequence 3, Appl
17	28.6	10.8	433	4 US-08-905-223-261	Sequence 261, App
18	27.8	10.5	248	4 US-09-007-005-32	Sequence 32, Appl
19	27.8	10.5	248	4 US-09-244-796-32	Sequence 32, Appl
20	27.8	10.5	277	4 US-09-007-005-3	Sequence 3, Appl
21	27.8	10.5	277	4 US-09-244-796-3	Sequence 3, Appl
22	27.8	10.5	614	4 US-09-328-111-212	Sequence 212, App
23	27.6	10.5	273	4 US-09-004-838-138	Sequence 138, App
24	27.6	10.5	2035	4 US-08-960-780-10	Sequence 10, Appl
25	27.6	10.5	2035	4 US-09-073-898-10	Sequence 10, Appl
26	27.6	10.5	3131	3 US-09-035-648-23	Sequence 23, Appl
27	27.6	10.5	3131	4 US-09-001-951-23	Sequence 23, Appl

C 28	27.6	10.5	3562	4 US-09-360-197-1	Sequence 1, Appl
C 29	27.6	10.5	3647	4 US-09-360-197-7	Sequence 7, Appl
C 30	27.6	10.5	18443	4 US-09-078-294-6	Sequence 6, Appl
C 31	27.4	10.4	420	4 US-09-328-111-413	Sequence 413, App
C 32	27.4	10.4	681	4 US-09-299-378-1	Sequence 1, Appl
C 33	27	10.2	933	6 US-09-344-12	Patent No. 5340934
34	27	10.2	2444	2 US-08-821-355A-2	Sequence 2, Appl
35	27	10.2	2444	2 US-09-003-687A-2	Sequence 2, Appl
36	27	10.2	2444	3 US-09-136-605-2	Sequence 2, Appl
37	27	10.2	11495	4 US-09-056-105-9	Sequence 9, Appl
38	26.8	10.2	1255	1 US-08-518-878B-38	Sequence 38, Appl
39	26.8	10.2	1255	1 US-08-294-522B-38	Sequence 38, Appl
40	26.8	10.2	1255	2 US-08-470-868A-38	Sequence 38, Appl
41	26.8	10.2	1545	1 US-08-400-275-17	Sequence 17, Appl
42	26.8	10.2	1596	2 US-08-807-861A-38	Sequence 38, Appl
43	26.8	10.2	1596	3 US-09-210-681-38	Sequence 38, Appl
44	26.8	10.2	1596	3 US-08-946-719A-38	Sequence 38, Appl
45	26.8	10.2	2125	4 US-09-303-639-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEFFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-Fls
US-08-232-463-14


```
Query Match 19.4% Score 51.2; DB 1; Length 7218;
Best Local Similarity 5.5%; Pred. No. 3.7e-08;
Matches 11; Conservative 128; Mismatches 61; Indels 0; Gaps 0;

QY 15 ttcccaatggttcggttcccaacttccaaagcccaacttgccttccatcaga 74
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1251 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1310
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 75 ggcctcgcccccagatctcgcttcctcaacttccaaacttccatcgcttlaa 134
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1311 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1370
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 135 cacaacaaactaaccttcgcttccttcctcgatccatccatcgctcccttaaacctcc 194
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1371 yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy 1430
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 195 ttacgatacaaatgcctcca 214
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1431 yyyggtaccaaattcttcta 1450

RESULT 2
US-09-356-952-9
; Sequence 9, Application US/09356952
; Patent No. 6117663
; GENERAL INFORMATION:
; APPLICANT: Borjask-Sjodin, Ann
; APPLICANT: Margalit, S. M.
; APPLICANT: Bor-Sogal, Dafna
; APPLICANT: Cole, Philip
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 600-1-228N
; CURRENT APPLICATION NUMBER: US/09/356,952
; EARLIER APPLICATION NUMBER: 60/093,631
; EARLIER FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 9
; LENGTH: 4002
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-356-952-9

Query Match 11.9% Score 31.4; DB 3; Length 4002;
Best Local Similarity 54.9%; Pred. No. 0.28;
Matches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 83 ccccgagattctcgcttcctcaacttcaaaccttctcatttcaacaaac 142
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3665 caccgtatgttttctcaagctcacacactacatctccaaacccccccttggtgacaa 3724
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 143 taacctatcgcttcctcccgatccatccatccatccctttaaactcct 195
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3725 gtgacacatgcaatgcttcttcccaaacagccctccccccttaacacact 3777

RESULT 3
US-08-658-136-2
; Sequence 2, Application US/08658136
; Patent No. 6071717
; GENERAL INFORMATION:
; APPLICANT: KLINGER, KATHERINE W
; APPLICANT: LANDES, GREGORY M
; APPLICANT: BURN, TIMOTHY C
; APPLICANT: CONNORS, TIMOTHY D
; APPLICANT: DACKOWSKI, WILLIAM
; APPLICANT: GERRINO, GREGORY
; APPLICANT: QIAN, FENG
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
```

```
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESS: GENZYME CORPORATION
STREET: ONE MOUNTAIN ROAD
CITY: FRAMINGHAM
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/658,136
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: LASSEN, ELIZABETH
REGISTRATION NUMBER: 31,845
REFERENCE/DOCKET NUMBER: GEN4-17.8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-872-8400
TELEFAX: 508-872-5415
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 53526 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-658-136-2

Query Match 11.8% Score 31.2; DB 3; Length 53526;
Best Local Similarity 47.4%; Pred. No. 1.2;
Matches 93; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 18 ccaatggttcggttcccaagcttcaagcgaatccatcgcttccttatagaag 77
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 35274 cccatgccccccttcctccctccctccctccctccctccctccctccctccctcc 35333
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 78 ctgcgccccgagattctcgcttcctcaacttcaaaccttctcatttcaac 137
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 35334 atccccctctctccctccctccctccctccctccctccctccctccctccctcc 35393
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 138 aaactaacctatcgcttcctcccgatccatccatccatccatccatcc 197
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 35394 ctctctccctccctccctccctccctccctccctccctccctccctccctcc 35453
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 198 cggatcaaatgctccc 213
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 35454 cctcccccactcctc 35469

RESULT 4
US-08-658-136-1
; Sequence 1, Application US/08658136
; Patent No. 6071717
; GENERAL INFORMATION:
; APPLICANT: KLINGER, KATHERINE W
; APPLICANT: LANDES, GREGORY M
; APPLICANT: BURN, TIMOTHY C
; APPLICANT: CONNORS, TIMOTHY D
; APPLICANT: DACKOWSKI, WILLIAM
; APPLICANT: GERRINO, GREGORY
; APPLICANT: QIAN, FENG
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
; NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESS: GENZYME CORPORATION
STREET: ONE MOUNTAIN ROAD
CITY: FRAMINGHAM
```


Oy 17 cccaatgctcgcgttccgaagctcaagaacccaacttcgtctctcatcagaagc 76
::: : : : : : : : : : : : :
Db 231 YCTATGCTTGTGYSYNINYSYNINYSYNINYSYNINYSYNINYSYNIN 172

Oy 77 gcttcgcccccagatctcgtctcctcacacctcaaacctcttcacatcgcttta 136

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 07:58:30 ; Search time 796.36 Seconds

(without alignments)
646,786 Million cell updates/sec

Title: US-09-300-482-619

Perfect score: 300
Sequence: 1 gattattgcaaacacgcgatt.....ttcagaatgtctgctgtg 300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 segs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	197.6	65.9	1836	21	AAAC42921
2	36.2	12.1	10820	20	AAAX20527
3	36	12.0	2375	21	AAAC42097
4	36	12.0	2614	21	AAAC36886
5	34.6	11.5	3436	23	ABLI17694
6	34.6	11.5	4605	23	ABLI23418
7	34	11.3	579	23	AAAS4595
8	33.2	11.1	7178	23	ABLO7679
9	33.2	11.1	12400	23	ABLO7678

10	32.6	10.9	222	22	AAAD19956	Human colon cancer
11	32.4	10.8	543	18	AAV75242	Staphylococcus aur
12	32.4	10.8	579	23	AAAS5056	Staphylococcus aur
13	31.6	10.5	11161	22	AAK82196	Human immune/haema
14	31.6	10.5	11934	22	AAK82197	Human immune/haema
15	31.6	10.5	32768	20	AAK13336	Enterococcus faeca
16	30.8	10.3	579	20	AAK87836	S. aureus phosphat
17	30.8	10.3	1004	20	AAK13479	Human musculoskele
18	30.8	10.3	1056	22	AAI36027	Human immune/haema
19	30.6	10.2	812	22	AAK56634	Streptococcus pneu
20	30.6	10.2	1051	19	AAV27386	Streptococcus pneu
21	30.6	10.2	13206	19	AAV52166	Staphylococcus aur
22	30.6	10.2	22243	18	AAV74475	Staphylococcus aur
23	30.4	10.1	163319	21	AAE22306	Arabidopsis thalia
24	30	10.0	299	23	AAAS50154	Staphylococcus aur
25	30	10.0	450	22	ABA20394	Human nervous syst
26	30	10.0	450	22	ABA20395	Human nervous syst
27	30	10.0	450	22	ABA20397	Human nervous syst
28	30	10.0	450	22	ABA20400	Human nervous syst
29	30	10.0	1725	23	AAAS55216	Staphylococcus aur
30	30	10.0	6712	24	ABL32690	Human immune syste
31	29.8	9.9	1062	20	AAZ39041	Human FLICE-like 1
32	29.8	9.9	1190	19	AAV61936	Human CFTRP-S DNA.
33	29.8	9.9	1308	20	AAZ55130	Human MACH related
34	29.8	9.9	1373	19	AAV52869	Human G1 protein i
35	29.8	9.9	1683	21	AAZ99543	Human Insulin-resp
36	29.8	9.9	1750	20	AAV74136	Human FLAME-1 cDNA
37	29.8	9.9	1996	20	AAZ5508	Human MACH related
38	29.8	9.9	2034	19	AAV44806	I-FLICE-1 coding s
39	29.8	9.9	2040	21	AAZ48769	Human CLARP coding
40	29.8	9.9	2045	19	AAV50436	Human Casper gene.
41	29.8	9.9	2064	21	AAZ99539	Human Insulin-resp
42	29.8	9.9	2084	21	AAZ56987	Human Insulin-resp
43	29.8	9.9	2138	19	AAV24137	Human FLICE-like 1
44	29.8	9.9	2143	19	AAV61937	Human CFTRP-L DNA.
45	29.8	9.9	2143	20	AAZ39040	Human FLICE-like 1

ALIGNMENTS

RESULT	1
ID	AAAC42921
AC	AAAC42921 standard; DNA; 1836 BP.
XX	
XX	AAAC42921;
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 37360.
XX	
KW	Hybridisation assay; genetic mapping; gene expression control
KW	protein identification; signal transduction pathway;
KW	metabolic pathway; promoter; termination sequence; ss.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0132866.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138447.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151309.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155146.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR	23-APR-1999;	990S-0130510.
PR	23-APR-1999;	990S-0130891.
PR	28-APR-1999;	990S-0131449.
PR	30-APR-1999;	990S-0132048.
PR	04-MAY-1999;	990S-0132407.
PR	05-MAY-1999;	990S-0132484.
PR	06-MAY-1999;	990S-0132485.
PR	06-MAY-1999;	990S-0132486.
PR	07-MAY-1999;	990S-0132487.
PR	11-MAY-1999;	990S-0132863.
PR	14-MAY-1999;	990S-0134216.
PR	14-MAY-1999;	990S-0134219.
PR	14-MAY-1999;	990S-0134221.
PR	18-MAY-1999;	990S-0134370.
PR	18-MAY-1999;	990S-0134376.
PR	19-MAY-1999;	990S-0134941.
PR	21-MAY-1999;	990S-0135124.
PR	24-MAY-1999;	990S-0135353.
PR	25-MAY-1999;	990S-0136021.
PR	27-MAY-1999;	990S-0136392.
PR	28-MAY-1999;	990S-0136782.
PR	03-JUN-1999;	990S-0137222.
PR	03-JUN-1999;	990S-0137228.
PR	04-JUN-1999;	990S-0137502.
PR	07-JUN-1999;	990S-0137724.
PR	08-JUN-1999;	990S-0138094.
PR	10-JUN-1999;	990S-0138540.
PR	10-JUN-1999;	990S-0138847.
PR	14-JUN-1999;	990S-0139119.
PR	16-JUN-1999;	990S-0139452.
PR	16-JUN-1999;	990S-0139453.
PR	17-JUN-1999;	990S-0139492.
PR	18-JUN-1999;	990S-0139454.
PR	18-JUN-1999;	990S-0139455.
PR	18-JUN-1999;	990S-0139456.
PR	18-JUN-1999;	990S-0139457.
PR	18-JUN-1999;	990S-0139458.
PR	18-JUN-1999;	990S-0139459.
PR	18-JUN-1999;	990S-0139460.
PR	18-JUN-1999;	990S-0139461.
PR	18-JUN-1999;	990S-0139462.
PR	18-JUN-1999;	990S-0139463.
PR	18-JUN-1999;	990S-0139750.
PR	18-JUN-1999;	990S-0139763.
PR	21-JUN-1999;	990S-0139817.
PR	22-JUN-1999;	990S-0139859.
PR	23-JUN-1999;	990S-0140053.
PR	23-JUN-1999;	990S-0140054.
PR	24-JUN-1999;	990S-0140695.
PR	28-JUN-1999;	990S-0140823.
PR	29-JUN-1999;	990S-0140991.
PR	30-JUN-1999;	990S-0141287.
PR	01-JUL-1999;	990S-0141842.
PR	01-JUL-1999;	990S-0142154.
PR	02-JUL-1999;	990S-0142055.
PR	06-JUL-1999;	990S-0142350.
PR	09-JUL-1999;	990S-0142803.
PR	09-JUL-1999;	990S-0142920.
PR	12-JUL-1999;	990S-0142977.
PR	13-JUL-1999;	990S-0143542.
PR	14-JUL-1999;	990S-0143624.
PR	15-JUL-1999;	990S-0144005.
PR	16-JUL-1999;	990S-0144085.
PR	16-JUL-1999;	990S-0144086.
PR	19-JUL-1999;	990S-0144325.
PR	19-JUL-1999;	990S-0144331.
PR	19-JUL-1999;	990S-0144332.
PR	19-JUL-1999;	990S-0144333.
PR	19-JUL-1999;	990S-0144334.
PR	19-JUL-1999;	990S-0144335.
PR	20-JUL-1999;	990S-0144352.
PR	23-APR-1999;	990S-0144632.
PR	20-JUL-1999;	990S-0144884.
PR	21-JUL-1999;	990S-0144814.
PR	21-JUL-1999;	990S-0145088.
PR	22-JUL-1999;	990S-0145087.
PR	22-JUL-1999;	990S-0145089.
PR	22-JUL-1999;	990S-0145091.
PR	22-JUL-1999;	990S-0145092.
PR	23-JUL-1999;	990S-0145119.
PR	23-JUL-1999;	990S-0145145.
PR	23-JUL-1999;	990S-0145218.
PR	26-JUL-1999;	990S-0145224.
PR	27-JUL-1999;	990S-0145913.
PR	27-JUL-1999;	990S-0145918.
PR	27-JUL-1999;	990S-0145919.
PR	28-JUL-1999;	990S-0145951.
PR	02-AUG-1999;	990S-0146386.
PR	02-AUG-1999;	990S-0146388.
PR	02-AUG-1999;	990S-0146389.
PR	03-AUG-1999;	990S-0147038.
PR	04-AUG-1999;	990S-0147204.
PR	05-AUG-1999;	990S-0147302.
PR	05-AUG-1999;	990S-0147309.
PR	06-AUG-1999;	990S-0147309.
PR	06-AUG-1999;	990S-0147413.
PR	09-AUG-1999;	990S-0147935.
PR	10-AUG-1999;	990S-0148171.
PR	11-AUG-1999;	990S-0148341.
PR	12-AUG-1999;	990S-0

PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI: 2001-656860/75.
XX
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS
CC Claim 1: SEQ ID NO 4555; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3436 BP; 836 A; 970 C; 681 G; 949 T; 0 other;

Query Match 11.5%; Score 34.6; DB 23; Length 3436;
Best Local Similarity 49.2%; Pred. No. 0.65;
Matches 91; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 94 agagcgaggagcaacctgaaccgcaatgctctactagaaglacagaagcctcagag 153
DB 3099 ATAGCATTTGGAAATCATTCACCTGCTATTTTATAGCCGGAATACAAAGCTCTTCTGTG 3040
QY 154 atgcggggctgaatctctcgaacaagggtgtgcaattactcaagaanaattctctgtg 213
DB 3039 ATACGATATTAATGCTATACATTGAGGCTTTTATATTAATAAGATAAATTTTAAATA 2980
QY 214 ataacactgctagaatagaaggagtgtagctcggttctcctagtgttgatggtgtg 273
DB 2979 GCTGCGCTTACAAATAATAGATCATATATTCACGTGGGTTTTTAACTCGTAGTGTGCGG 2920
QY 274 gtagg 278
DB 2919 CGAGG 2915

RESULT 6
ID ABL23418 standard; DNA; 4605 BP.
AC ABL23418;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 21727.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.

XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI: 2001-656860/75.
XX
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS
CC Claim 1: SEQ ID NO 21727; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB572072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4605 BP; 1222 A; 787 C; 1319 G; 1277 T; 0 other;

Query Match 11.5%; Score 34.6; DB 23; Length 4605;
Best Local Similarity 49.2%; Pred. No. 0.73;
Matches 91; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
QY 94 agagcgaggagcaacctgaaccgcaatgctctactagaaglacagaagcctcagag 153
DB 534 ATAGCATTTGGAAATCATTCACCTGCTATTTTATAGCCGGAATACAAAGCTCTTCTGTG 475
QY 154 atgcggggctgaatctctcgaacaagggtgtgcaattactcaagaanaattctctgtg 213
DB 474 ATACGATATTAATGCTATACATTGAGGCTTTTATATTAATAAGATAAATTTTAAATA 415
QY 214 ataacactgctagaatagaaggagtgtagctcggttctcctagtgttgatggtgtg 273
DB 414 GCTGCGCTTACAAATAATAGATCATATATTCACGTGGGTTTTTAACTCGTAGTGTGCGG 355
QY 274 gtagg 278
DB 354 CGAGG 350

RESULT 7
ID AAS54595 standard; DNA; 579 BP.
AC AAS54595;
XX
DT 13-FEB-2002 (first entry)
XX
DE Staphylococcus aureus DNA for cellular proliferation protein #907.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-20727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.

XX Polynucleotide(s) and proteins derived from *Staphylococcus aureus*
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
XX

PS Claim 1; Page 1752-1753; 3271pp; English.

XX This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the *S.aureus* DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against *S.aureus* infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC *S.aureus* in a sample. *S.aureus* is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the *S.aureus* DNA sequences contained on the
CC computer readable medium.
XX

SO Sequence 543 BP; 178 A; 100 C; 53 G; 152 T; 60 other;

Query Match 10.8%; Score 32.4; DB 18; Length 543;
Best Local Similarity 50.6%; Pred No.1,7;
Matches 78; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

OY 41 aattgtctaactaggaacctggaactctctgttaattgcatcttaagaagcgg 100
DB 207 AATTGTACTTGTGCAACTAGGACTAACAATTCTGTAGCAATCATTTATTATGGCAG 148
OY 101 aggcacacctgaaacccgcaatgctctactagaagtagaagaagcctcaagatgcgg 160
DB 147 AGAATTTGCCGTACTGGTTTACTGTACTCAAAATTGAACAAGGATTCGTAAGTGACG 88
OY 161 gctgcaattctcgaaacagggtgttgcaattact 194
DB 87 TGGTCATTTAGTAAATTAAACAGCAGTTACT 54

RESULT 12
AAS55056

ID AAS55056 standard; DNA; 579 BP.

AC AAS55056;

DT 13-FEB-2002 (first entry)

DE *Staphylococcus aureus* DNA for cellular proliferation protein #1368.

KW Antisense; ds; prokaryotic cellular proliferation gene;
KW antibiotic; antibacterial; drug design.

OS *Staphylococcus aureus*.

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001MO-0509180.

XX 21-MAR-2000; 2000US-191078P.

XX 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 24-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyaskind JW, Wall D, Trawack JD, Carr GJ;
PI Yamamoto RT, Xu HH;

DR WPI; 2001-611495/70.
P-PSDB; AAU37197.

XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX

PS Claim 27; Seq ID No 8693; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acids which are required to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

SO Sequence 579 BP; 176 A; 63 C; 115 G; 225 T; 0 other;

Query Match 10.8%; Score 32.4; DB 23; Length 579;
Best Local Similarity 50.6%; Pred No.1,7;
Matches 78; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

OY 41 aattgtctaactaggaacctggaactctctgttaattgcatcttaagaagcgg 100
DB 270 aattgtactgtgcaactaggaactaactctgtgtagaagatcatatattatgccag 329
OY 101 aggcacacctgaaacccgcaatgctctactagaagtagaagaagcctcaagatgcgg 160
DB 330 agaattgcccgaactcgttctactactacaattgaaacaaagatcgtgaagtcgacg 389
OY 161 gctgcaattctcgaaacagggtgttgcaattact 194
DB 390 tgtcgaattaggttaataataaacagcagttact 423

RESULT 13
AAK82196/C

ID AAK82196 standard; DNA; 11161 BP.

AC AAK82196;

DT 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:37008.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX cytostatic; gene therapy; vaccine; metastasis; ds.

OS *Homo sapiens*.

PN WO200157182-A2.

PD 09-AUG-2001.

XX. 17-JAN-2001; 2001WO-US01354.
PF
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184654.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234937.
PR 25-SEP-2000; 2000US-0234938.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236357.
PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251858.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254059.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX

PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 37008; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 11161 BP; 3362 A; 1862 C; 1811 G; 4126 T; 0 other;

Query Match 10.5%; Score 31.6; DB 22; Length 11161;
Best Local Similarity 51.4%; Pred. No. 9.7;
Matches 73; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 60 gaactggcaactactctgttaattgtcattcttaagagcgagcgacccctgaacccgc 119
DB 7181 GCACCTTAATAATGACATATTGTATGCACAGACATCTTGGCCAAATTAACCTGGAACCTAG 7122
QY 120 aatgtgtctctcgaagtagacagaagccttcagagatcgcggtcgaattctcgaacag 179
DB 7121 AAGGGCTCTCCCTGAATCTTAAAGAACATGACGTAGATGATGCATATATAAGAACAAA 7062
QY 180 ggtgtgcaattaccacgaaga 201
DB 7061 TCAGAAAGATTAATAAAAAAAA 7040

RESULT 14
AAK82197/c
ID AAK82197 standard; DNA; 11934 BP.
AC XX
XX AAK82197;
DT 07-NOV-2001 (first entry)
XX
DE Human immune/hematopoietic antigen genomic sequence SEQ ID NO:37009.
XX
KW Human; immune; haematopoietic; immune/hematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
PD 09-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.

PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241786.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
XX		
DR	WPI; 2001-483426/52.	
XX		
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and	
PT	metastasis -	
XX		
PS	Disclosure; SEQ ID NO 37009; 3071pp + Sequence Listing; English.	
XX		
CC	AAK4951 to AAK64702 encode the human immune/hematopoietic antigen (II)	
CC	amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic	
CC	activity, and can be used in gene therapy and vaccine production. (II)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (II) expression. For	

CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK4950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 11934 BP; 3616 A; 1963 C; 1946 G; 4409 T; 0 other;

[illegible]

```

RESULT_15
AAAX13336
ID AAAX13336 standard; DNA; 32768 BP.
XX
XX AC AAAX13336;
XX
XX DT 19-MAR-1999 (first entry)
XX
DE Enterococcus faecalis genome contig SEQ ID NO:399.
XX
XX KW Enterococcus faecalis; contig; detection; Enterococcal infection;
XX KW vaccine; attenuation; computer readable medium; ds.
XX
XX OS Enterococcus faecalis.
XX
XX PN W09850555-A2.
XX
XX PD 12-NOV-1998.
XX
XX PF 04-MAY-1998; 98WO-US08985.
XX
XX PR 14-NOV-1997; 97US-0066009.
XX PR 06-MAY-1997; 97US-0044031.
XX PR 16-MAY-1997; 97US-0046555.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Barash SC, Dillon PJ, Kunsch CA;
XX
XX DR WPI: 1999-045171/04.
XX
XX PT New isolated Enterococcus faecalis polynucleotides and polypeptides
XX PT - used to develop products for the detection of Enterococcus and for
XX PT use in vaccines for prevention or attenuation of Enterococcus
XX PT infection.
XX
XX PS Claim 1; Page 1592-1609; 2084pp; English.
XX
CC A computer readable medium has been developed which has recorded on it
CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.

```


CC AAX12938 to AAX13919 represent these nucleotide sequences which are
CC primary nucleotide sequences, also known as contigs. The computer-based
CC system can identify fragments of the Enterococcus faecalis genome with
CC commercial importance. The products can be used to detect the presence
CC of Enterococcus faecalis in samples. They can also be used for
CC diagnosing Enterococcal infection in an animal and monitoring
CC progression of disease, and for identifying agents which can be used to
CC modulate the growth or pathogenicity of Enterococcus faecalis, or
CC another related organism, in vivo or in vitro. In particular the
CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
CC can be used in vaccines to prevent or attenuate an Enterococcal
CC infection.

CC
XX
SQ Sequence 32768 BP; 10797 A; 5960 C; 7083 G; 8921 T; 7 other;

Query Match 10.5%; Score 31.6; DB 20; Length 32768;
Best Local Similarity 51.4%; Pred. No. 15;
Matches 73; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 52 taggacctgaactggcaactactctgttaattgtcatcttctaagagcgaggcacactg 111
Db 4846 ttggaccgcccgaaggtaataagttcatatagtgttcttgaataacagcacacta 4905
QY 112 aaaccgcaatggtctactagaagaagccttcagagatgcygggtcgaattct 171
Db 4906 aattcatcaattcatatataagtttttgaatgagtaacgaacagacaacaaatca 4965
QY 172 cgaacaagggtgttgcattac 193
Db 4966 aggtcgtggtgttccatctcc 4987

Search completed: July 3, 2002, 07:58:41
Job time: 24476 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 11:18:14 ; Search time 10825.5 Seconds

(without alignments)
599.556 Million cell updates/sec

Title: US-09-300-482-619

Perfect score: 300
Sequence: 1 gattattgcacacccgat.....ttcagaatgtctgctgtg 300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA Main: *
1: /cgn2_6/ptodata/2/pna/PCrUS_COMB.seq: *
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq: *
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq: *
4: /cgn2_6/ptodata/2/pna/US08_COMB.seq: *
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq: *
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq: *
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq: *
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq: *
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq: *
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq: *
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq: *
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq: *
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq: *
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq: *
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq: *
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq: *
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq: *
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq: *
19: /cgn2_6/ptodata/2/pna/US095_COMB.seq: *
20: /cgn2_6/ptodata/2/pna/US095b_COMB.seq: *
21: /cgn2_6/ptodata/2/pna/US095c_COMB.seq: *
22: /cgn2_6/ptodata/2/pna/US095d_COMB.seq: *
23: /cgn2_6/ptodata/2/pna/US096a_COMB.seq: *
24: /cgn2_6/ptodata/2/pna/US096b_COMB.seq: *
25: /cgn2_6/ptodata/2/pna/US096c_COMB.seq: *
26: /cgn2_6/ptodata/2/pna/US096d_COMB.seq: *
27: /cgn2_6/ptodata/2/pna/US096e_COMB.seq: *
28: /cgn2_6/ptodata/2/pna/US097a_COMB.seq: *
29: /cgn2_6/ptodata/2/pna/US097b_COMB.seq: *
30: /cgn2_6/ptodata/2/pna/US097c_COMB.seq: *
31: /cgn2_6/ptodata/2/pna/US098a_COMB.seq: *
32: /cgn2_6/ptodata/2/pna/US098b_COMB.seq: *
33: /cgn2_6/ptodata/2/pna/US098c_COMB.seq: *
34: /cgn2_6/ptodata/2/pna/US099a_COMB.seq: *
35: /cgn2_6/ptodata/2/pna/US099b_COMB.seq: *
36: /cgn2_6/ptodata/2/pna/US099c_COMB.seq: *
37: /cgn2_6/ptodata/2/pna/US100_COMB.seq: *
38: /cgn2_6/ptodata/2/pna/US101_COMB.seq: *
39: /cgn2_6/ptodata/2/pna/US600_COMB.seq: *
40: /cgn2_6/ptodata/2/pna/US6001_COMB.seq: *
41: /cgn2_6/ptodata/2/pna/US6002_COMB.seq: *
42: /cgn2_6/ptodata/2/pna/US6003_COMB.seq: *
43: /cgn2_6/ptodata/2/pna/US6004_COMB.seq: *

Result				SUMMARIES			
No.	Score	Match	Query Length	ID	Description		
1	300	100.0	300	16	US-09-237-183A-276	Sequence 276, App	
2	300	100.0	300	17	US-09-300-482-619	Sequence 619, App	
3	300	100.0	300	17	US-09-304-517A-10554	Sequence 10554, A	
4	300	100.0	300	17	US-09-371-146A-10554	Sequence 10554, A	
5	300	100.0	300	36	US-09-985-678-10554	Sequence 10554, A	
6	296.8	98.9	2199	33	US-09-865-439A-11076	Sequence 11076, A	
7	295.2	98.4	434	27	US-09-696-664A-7280	Sequence 7280, App	
8	289.8	96.3	1850	25	US-09-654-617-267132	Sequence 267132, A	
9	288.8	96.3	1850	25	US-09-654-617-267132	Sequence 267132, A	
10	288.8	96.3	1850	27	US-09-684-016-51737	Sequence 51737, A	
11	285.6	95.2	590	33	US-09-865-419A-51737	Sequence 388605, A	
12	223.8	74.6	1056	25	US-09-684-016-388605	Sequence 388605, A	
13	223.8	74.6	1056	27	US-09-684-016-388605	Sequence 388605, A	
14	218.4	72.8	509	32	US-09-849-529A-21745	Sequence 21745, A	
15	218.4	72.8	509	38	US-09-196-868-21732	Sequence 21732, A	
16	208.8	69.6	397	33	US-09-874-708A-57882	Sequence 57882, A	
17	208.8	69.6	397	60	US-60-211-750-56895	Sequence 56895, A	
18	208.8	69.6	1156	28	US-09-705-926-2660	Sequence 2660, App	
19	208.8	69.6	1368	25	US-09-654-617-3903	Sequence 3903, App	
20	208.8	69.6	1368	27	US-09-684-016-46195	Sequence 46195, A	
21	208.8	69.6	2218	33	US-09-874-708A-46195	Sequence 46195, A	
22	203.4	67.8	362	26	US-09-669-817A-33549	Sequence 33549, A	
23	197.6	65.9	1836	28	US-09-708-427-10377	Sequence 10377, A	
24	197.6	65.9	1836	35	US-09-938-842A-12322	Sequence 12322, App	
25	194.8	64.9	470	25	US-09-654-617-405674	Sequence 405674, A	
26	194.8	64.9	470	27	US-09-684-016-405674	Sequence 405674, A	
27	194.8	64.9	470	58	US-60-197-872-53225	Sequence 53225, A	
28	194.8	64.9	2413	33	US-09-867-716-19334	Sequence 19334, A	
29	192.8	64.3	343	64	US-60-253-456-3653	Sequence 3653, App	
30	190	63.3	593	64	US-60-253-456-28250	Sequence 28250, A	
31	184	61.3	411	25	US-09-654-617-234017	Sequence 234017, A	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.


```
32      184      61.3      411      26      US-09-666-355A-23565      Sequence 23565, A
33      184      61.3      411      27      US-09-664-016-234017      Sequence 234017,
34      175.4      58.5      481      33      US-09-874-708A-5054      Sequence 5054, Ap
35      175.4      58.5      481      60      US-09-211-750-4911      Sequence 4911, Ap
36      174.4      58.1      488      33      US-09-874-708A-520      Sequence 520, App
37      174.4      58.1      488      60      US-09-211-750-520      Sequence 520, App
38      159.8      53.3      235      16      US-09-237-183A-285      Sequence 285, App
39      159.8      53.3      235      17      US-09-300-482-628      Sequence 628, App
40      159.8      53.3      235      17      US-09-304-517A-22437      Sequence 22437, A
41      159.8      53.3      235      17      US-09-371-146A-22437      Sequence 22437, A
42      159.8      53.3      235      36      US-09-985-678-22437      Sequence 22437, A
43      159.8      53.3      248      16      US-09-288-328A-664      Sequence 664, App
44      159.8      53.3      248      47      US-09-085-147-664      Sequence 664, App
45      145      46.3      248      16      US-09-237-183A-282      Sequence 282, App
```

ALIGNMENTS

```
RESULT 1
US-09-237-183A-276
; Sequence 276, Application US/09237183A
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Fisher, Dane K.
; APPLICANT: Liu, Jindong
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With The
; FILE REFERENCE: 38-21(15089)B
; CURRENT APPLICATION NUMBER: US/09/237,183A
; PRIOR FILING DATE: 1999-01-26
; PRIOR APPLICATION NUMBER: US 60/067,000
; NUMBER OF SEQ ID NOS: 2814
; SEQ ID NO 276
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Zea mays
US-09-237-183A-276
```

```
Query Match      100.0%; Score 300; DB 16; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.6e-87;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 gattattgaacacccgatactctgtgtggaatgatacaaatgtctcaactagagactg 60
    |||
Db 1 gattattgaacacccgatactctgtgtggaatgatacaaatgtctcaactagagactg 60
QY 61 aactggcaactactctgtgaatgtatcttctaagagcgagggcacactgaaacccgca 120
    |||
Db 61 aactggcaactactctgtgaatgtatcttctaagagcgagggcacactgaaacccgca 120
QY 121 atggtctactagaagtaacagaagccttcagagatcgaggcgacgaatctctgaaacag 180
    |||
Db 121 atggtctactagaagtaacagaagccttcagagatcgaggcgacgaatctctgaaacag 180
QY 181 gtgttgaattactcaagaataatctctgttgaatacaactctagaaatagagggatggt 240
    |||
Db 181 gtgttgaattactcaagaataatctctgttgaatacaactctagaaatagagggatggt 240
QY 241 tagctcggttccatactgttgaatgtgtgtgtgtagagcctcagaagaatgtctgtg 300
    |||
Db 241 tagctcggttccatactgttgaatgtgtgtgtgtagagcctcagaagaatgtctgtg 300
```

```
RESULT 2
US-09-300-482-619
; Sequence 619, Application US/09300482
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Liu, Jindong
; APPLICANT: Peschke, Virginia M.
```

```
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With The
; FILE REFERENCE: 04983.0031.us01/38-21(15365)B
; CURRENT APPLICATION NUMBER: US/09/300,482
; CURRENT FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 699
; SEQ ID NO 619
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Zea mays
US-09-300-482-619
```

```
Query Match      100.0%; Score 300; DB 17; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.6e-87;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 gattattgaacacccgatactctgtgtggaatgatacaaatgtctcaactagagactg 60
    |||
Db 1 gattattgaacacccgatactctgtgtggaatgatacaaatgtctcaactagagactg 60
QY 61 aactggcaactactctgtgaatgtatcttctaagagcgagggcacactgaaacccgca 120
    |||
Db 61 aactggcaactactctgtgaatgtatcttctaagagcgagggcacactgaaacccgca 120
QY 121 atggtctactagaagtaacagaagccttcagagatcgaggcgacgaatctctgaaacag 180
    |||
Db 121 atggtctactagaagtaacagaagccttcagagatcgaggcgacgaatctctgaaacag 180
QY 181 gtgttgaattactcaagaataatctctgttgaatacaactctagaaatagagggatggt 240
    |||
Db 181 gtgttgaattactcaagaataatctctgttgaatacaactctagaaatagagggatggt 240
QY 241 tagctcggttccatactgttgaatgtgtgtgtgtagagcctcagaagaatgtctgtg 300
    |||
Db 241 tagctcggttccatactgttgaatgtgtgtgtgtagagcctcagaagaatgtctgtg 300
```

```
RESULT 3
US-09-304-517A-10554
; Sequence 10554, Application US/09304517A
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Liu, Jindong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)B
; CURRENT APPLICATION NUMBER: US/09/304,517A
; CURRENT FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 10554
; LENGTH: 300
; TYPE: DNA
; ORGANISM: Zea mays
US-09-304-517A-10554
```

```
Query Match      100.0%; Score 300; DB 17; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.6e-87;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gattattgaacacccgatactctgtgtggaatgatacaaatgtctcaactagagactg 60
    |||
Db 1 gattattgaacacccgatactctgtgtggaatgatacaaatgtctcaactagagactg 60
QY 61 aactggcaactactctgtgaatgtatcttctaagagcgagggcacactgaaacccgca 120
    |||
Db 61 aactggcaactactctgtgaatgtatcttctaagagcgagggcacactgaaacccgca 120
QY 121 atggtctactagaagtaacagaagccttcagagatcgaggcgacgaatctctgaaacag 180
    |||
Db 121 atggtctactagaagtaacagaagccttcagagatcgaggcgacgaatctctgaaacag 180
QY 181 gtgttgaattactcaagaataatctctgttgaatacaactctagaaatagagggatggt 240
    |||
```


	+
Db	181 ggtgcacattaccacaagaataatctctgtttgataaacacgcctagaataagaggaatgct 240
Oy	241 tagctcggattccctacatgttttattgggttggtgtagacttcgaatagtctgcgttg 300
Db	241 tagctcggattccctacatgttttattgggttggtgtagacttcgaatagtctgcgttg 300

RESULT 4
US-09-371-146A-10554

```

: Sequence 10554, Application US/09371146A
: GENERAL INFORMATION:
: APPLICANT: Cheikh, Nordine
: APPLICANT: Liu, Jiongong
: TITLE OF INVENTION: ANNOTATED PLANT GENES
: FILE REFERENCE: 38-21(159)7C
: CURRENT APPLICATION NUMBER: US/09/371,146A
: CURRENT FILING DATE: 1999-08-09
: PRIOR APPLICATION NUMBER: US 09/304,517
: PRIOR FILING DATE: 1999-05-06
: NUMBER OF SEQ ID NOS: 294310
: SEQ ID NO 10554
: LENGTH: 300
: TYPE: DNA
: ORGANISM: Zea mays
: US-09-371-146A-10554

```

Query Match	100.0%;	Score 300;	DB 17;	Length 300;
Best Local Similarity	100.0%;	Pred. No. 2.6e-87;		
Matches 300;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Oy	1	gatttatatgaaacaacccgattctctggtgattgtatctcaaatgtcccaacttgacctg	60
Db	1	gatttatatgaaacaacccgattctctggtgattgtatctcaaatgtcccaacttgacctg	60
Oy	61	aacttgcacactactctctgttaattgtcatcttcaagagcgagacacccgaaacccga	120
Db	61	aacttgcacactactctctgttaattgtcatcttcaagagcgagacacccgaaacccga	120
Oy	121	atggttctaactgaaggtacagaagcccttcagagatgtgaggcgctgcgaattctcgaaagaag	180
Db	121	atggttctaactgaaggtacagaagcccttcagagatgtgaggcgctgcgaattctcgaaagaag	180
Oy	181	gtgttgcacattaccacaagaataatctctgtttgtgataaacactgctagagaataggatggt	240
Db	181	gtgttgcacattaccacaagaataatctctgtttgtgataaacactgctagagaataggatggt	240
Oy	241	tagctcgtgttctccatgatttggattggtgtgtgtgtagacttcagaatgtctcgtctgtg	300
Db	241	tagctcgtgttctccatgatttggattggtgtgtgtgtgtagacttcagaatgtctcgtctgtg	300

```

RESULTS 5
US-09-985-678-10554
: Sequence 10554, Application US/09985678
: GENERAL INFORMATION:
: APPLICANT: Chelrh, Norgine
: APPLICANT: Liu, Jingdong
: TITLE OF INVENTION: Annotated Plant Genes
: FILE REFERENCE: 16517, 255/38-11(15097)F
: CURRENT APPLICATION NUMBER: US/09/985, 678
: CURRENT FILING DATE: 2001-11-05
: PRIOR APPLICATION NUMBER: US 09/304,517
: PRIOR FILING DATE: 1999-03-06
: NUMBER OF SEQ ID NOS: 295529
: SEQ ID NO 10554
: LENGTH: 300
: TYPE: DNA
: ORGANISM: Zea mays
US-09-985-678-10554

```

Query Match	100.0%;	Score 300;	DB 36;	Length 300;
Best Local Similarity	100.0%;	Pred. No. 2.6e-87;		
Matches 300;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

1 gattattgacaaacccgatacctctggtggaattgacatcaaatgtctcaactagagacctg 60
|||||
1 gattattgacaaacccgatacctctggtggaattgacatcaaatgtctcaactagagacctg 60

61 aacggcgcaactactctgttaattgtcatcttctaagagcgagcacacactgaaaccgca 120
|||||
61 aacggcgcaactactctgttaattgtcatcttctaagagcgagcacacactgaaaccgca 120

121 atgtctactagaagtaacagaaagccttcagagatgcgggctgtcaattctcgaacagg 180
|||||
121 atgtctactagaagtaacagaaagccttcagagatgcgggctgtcaattctcgaacagg 180

181 gtttcgaacttactcaagaanaattctctctgttgatacaactcttagaatagaggatgt 240
 |||||
 181 gtgtttgcaattactcaagaanaattctctctgttgatacaactcttagaatagaggatgt 240

241 tagctcgtttccctatgatttgatgggttcgtgtgtagagacttcagaatgctctgtgttg 300

|||||

241 tagctcgtttccctatgatttgatgggttcgtgtgtagagacttcagaatgctctgtgttg 300

RESULT 6
IS-09-865-439A-11076/c

Sequence 11076, Application US/098655439A
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
APPLICANT: Hardeman, Kristine J.
APPLICANT: La Rosa, Thomas J.
TITLE OR INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
TITLE OR INVENTION: Plants
FILE REFERENCE: 38-21(51936)B
CURRENT APPLICATION NUMBER: US/09/865,439A
CURRENT FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: US 60/207,458
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ. ID NOS: 119126
SEQ. ID NO 11076
LENGTH: 2199
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3205-077-A12.FLI
US-09-865-439A-11076

Query Match	98.9%;	Score 296.8;	DB 33;	Length 2199;
Best Local Similarity	99.3%;	Pred. No. 6.6e-86;		
Matches 298; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

1	gattctctgacaacaacggatccctcgtcggaltgcatcacaattgctcaactagaaactg	60
1420	gattttttacacacacacccgattccctggctggattgattcatcaaatgtctcaactagaccctg	1361
61	aactggcaactactcctctgtaattgcatcttctaaagcggagcacaactggaaacccgca	120
1360	aaCTGGCAACTCTCTTGTAAATTGTCATTTTAAAGACGGAGGACACCTGAAACCCGCA	1301
121	atgcttactaagaaglacagaagagccttcagagaatgcggggcgtcgaattctcgaaacag	180
1300	ATGCTTACTAAGAAGACAGAAAGCCTTCACAGATGCAGCGGCTGCAATTCTGAAACAGG	1241
181	gttttcaactactcaaaaaaattctcgtttggatgaactgctaaattgaagggaatg	240
1240	GTTTTCAATTACTCAAGAAATTTCTCTGTTGGATTAACACTGCTTAAGAAATGAGGAGTGT	1181
241	taagctcgattctcctatgtttgatttggatctggatgtagaacttcaagaatgctctgctg	300
1180	TAACTCGATTCTCTATGTTTGAATTGGATGGGTGTAGACCTTACGAAGAATGTACGCTGGG	1121

RESULT 7

```
US-09-696-664A-7280
; Sequence 7280, Application US/09696664A
; GENERAL INFORMATION:
; APPLICANT: Abed, Mark S.
; APPLICANT: Andersen, Scott E.
; APPLICANT: Dubois, Patrice
; APPLICANT: Mahadeo, Debbie A.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51721)B
; CURRENT APPLICATION NUMBER: US/09/696,664A
; PRIOR FILING DATE: 2000-10-25
; PRIOR FILING DATE: 1999-10-26
; NUMBER OF SEQ ID NOS: 17472
; SEQ ID NO 7280
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3279-041-P1-K1-F8
US-09-696-664A-7280
```

Query Match 98.4%; Score 295.2; DB 27; Length 434;
Best Local Similarity 99.0%; Pred. No. 1,1e-85;

Matches 297; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 gattattgacaaacacgcgtctggtggtatgatacacaattgctcaactaggactg 60
   |||||||
DB 100 gattattgacaaacacgcgtctggtggtatgatacacaattgctcaactaggactg 159
QY 61 aactggcaactactcttgaatgttcaattcttaagagcgagacacccgaaaccgca 120
   |||||||
DB 160 aactggcaactactcttgaatgttcaattcttaagagcgagacacccgaaaccgca 219
QY 121 atggtctactagaagtagaagaagccttcagagatgctgggctgcgaattctcgaaacagg 180
   |||||||
DB 220 atggtctactagaagtagaagaagccttcagagatgctgggctgcgaattctcgaaacagg 279
QY 181 gtgtgcaattactcaagaagaattctctgtgtgataaactgctagaaatagaggatggt 240
   |||||||
DB 280 gtgtgcaattactcaagaagaattctctgtgtgataaactgctagaaatagaggatggt 339
QY 241 tagctcggttccctatgttgaattggtgtgtgtagagactcaagaatgctgctgtg 300
   |||||||
DB 340 tagctcggttccctatgttgaattggtgtgtgtagagactcaagaatgctgctgtg 399
```

RESULT 8

```
US-60-161-619-7119
; Sequence 7119, Application US/60161619
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; APPLICANT: Andersen, Scott E.
; APPLICANT: Dubois, Patrice
; APPLICANT: Mahadeo, Debbie A.
; APPLICANT: Masucci, James D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(51721)A
; CURRENT APPLICATION NUMBER: US/60/161,619
; CURRENT FILING DATE: 1999-10-26
; NUMBER OF SEQ ID NOS: 7685
; SEQ ID NO 7119
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3279-041-P1-K1-F8
```

US-60-161-619-7119

Query Match 98.4%; Score 295.2; DB 55; Length 434;
Best Local Similarity 99.0%; Pred. No. 1,1e-85;

Matches 297; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 gattattgacaaacacgcgtctggtggtatgatacacaattgctcaactaggactg 60
   |||||||
DB 100 gattattgacaaacacgcgtctggtggtatgatacacaattgctcaactaggactg 159
QY 61 aactggcaactactcttgaatgttcaattcttaagagcgagacacccgaaaccgca 120
   |||||||
DB 160 aactggcaactactcttgaatgttcaattcttaagagcgagacacccgaaaccgca 219
QY 121 atggtctactagaagtagaagaagccttcagagatgctgggctgcgaattctcgaaacagg 180
   |||||||
DB 220 atggtctactagaagtagaagaagccttcagagatgctgggctgcgaattctcgaaacagg 279
QY 181 gtgtgcaattactcaagaagaattctctgtgtgataaactgctagaaatagaggatggt 240
   |||||||
DB 280 gtgtgcaattactcaagaagaattctctgtgtgataaactgctagaaatagaggatggt 339
QY 241 tagctcggttccctatgttgaattggtgtgtgtagagactcaagaatgctgctgtg 300
   |||||||
DB 340 tagctcggttccctatgttgaattggtgtgtgtagagactcaagaatgctgctgtg 399
```

RESULT 9

```
US-09-654-617-267132
; Sequence 267132, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 267132
; LENGTH: 1850
; TYPE: DNA
; ORGANISM: Zea mays
US-09-654-617-267132
```

Query Match 96.3%; Score 288.8; DB 25; Length 1850;
Best Local Similarity 97.7%; Pred. No. 2,6e-83;

Matches 293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY 1 gattattgacaaacacgcgtctggtggtatgatacacaattgctcaactaggactg 60
   |||||||
DB 430 gattattgacaaacacgcgtctggtggtatgatacacaattgctcaactaggactg 489
QY 61 aactggcaactactcttgaatgttcaattcttaagagcgagacacccgaaaccgca 120
   |||||||
DB 490 aactggcaactactcttgaatgttcaattcttaagagcgagacacccgaaaccgca 549
QY 121 atggtctactagaagtagaagaagccttcagagatgctgggctgcgaattctcgaaacagg 180
   |||||||
DB 550 atggtctactagaagtagaagaagccttcagagatgctgggctgcgaattctcgaaacagg 609
QY 181 gtgtgcaattactcaagaagaattctctgtgtgataaactgctagaaatagaggatggt 240
   |||||||
DB 610 gtgtgcaattactcaagaagaattctctgtgtgataaactgctagaaatagaggatggt 669
QY 241 tagctcggttccctatgttgaattggtgtgtgtagagactcaagaatgctgctgtg 300
   |||||||
DB 670 tagctcggttccctatgttgaattggtgtgtgtagagactcaagaatgctgctgtg 729
```

RESULT 10

US-09-684-016-267132

Sequence 267132, Application US/09684016
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jindong
TITLE OF INVENTION: Annotated Plant Genes
FILE REFERENCE: 38-21(15097)D
CURRENT APPLICATION NUMBER: US/09/684,016
CURRENT FILING DATE: 2000-10-10
PRIOR APPLICATION NUMBER: US 09/654,617
PRIOR FILING DATE: 2000-09-05
NUMBER OF SEQ ID NOS: 463173
SEQ ID NO 267132
LENGTH: 1850
TYPE: DNA
ORGANISM: Zea mays
US-09-684-016-267132

Query Match 96.3%; Score 288.8; DB 27; Length 1850;
Best Local Similarity 97.7%; Pred. No. 2.6e-83;
Matches 293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 gattattgacacacccgctctgctggtgatcatcaaatctgctcaactagagacctg 60
|||||
DB 430 gattattgacacacccgctctgctggtgatcatcaaatctgctcaactagagacctg 489
QY 61 aactggaactactctctgtaattgcatcttcaagaagcgagacacactgaaccgca 120
|||||
DB 490 aactggaactactctctgtaattgcatcttcaagaagcgagacacactgaaccgca 549
QY 121 atggtctactagaatagacagaagccttcagagatcggtggtgcaattctcgaaacag 180
|||||
DB 550 atggtctactagaatagacagaagccttcagagatcggtggtgcaattctcgaaacag 609
QY 181 gttgtgcaattcactcaagaanaattctctgttgatacaactctgataagagggatggt 240
|||||
DB 610 gttgtgcaattcactcaagaanaattctctgttgatacaactctgataagagggatggt 669
QY 241 tagctcggttccctatgtttgattggtgtgtgtagagacttcagaanaatgctgtgtg 300
|||||
DB 670 tagctcggttccctatgtttgattggtgtgtgtagagacttcagaanaatgctgtgtg 729

RESULT 11
US-09-865-419A-51737
Sequence 51737, Application US/09865419A
GENERAL INFORMATION:
APPLICANT: Conner, Timothy W.
APPLICANT: Wu, Kunsheng
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(51935)B
CURRENT APPLICATION NUMBER: US/09/865,419A
CURRENT FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: US 60/208,063
PRIOR FILING DATE: 2000-05-31
NUMBER OF SEQ ID NOS: 54020
SEQ ID NO 51737
LENGTH: 590
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(590)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: LIB3637-255-Q6-K6-D10
US-09-865-419A-51737

Query Match 95.2%; Score 285.6; DB 33; Length 590;
Best Local Similarity 97.0%; Pred. No. 1.8e-82;
Matches 291; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 gattattgacacacccgctctgctggtgatcatcaaatctgctcaactagagacctg 60
|||||
DB 81 gattattgacacacccgctctgctggtgatcatcaaatctgctcaactagagacctg 140
QY 61 aactggaactactctctgtaattgcatcttcaagaagcgagacacactgaaccgca 120
|||||
DB 141 aactggaactactctctgtaattgcatcttcaagaagcgagacacactgaaccgca 200
QY 121 atggtctactagaatagacagaagccttcagagatcggtggtgcaattctcgaaacag 180
|||||
DB 201 atggtctactagaatagacagaagccttcagagatcggtggtgcaattctcgaaacag 260
QY 181 gttgtgcaattcactcaagaanaattctctgttgatacaactctgataagagggatggt 240
|||||
DB 261 gttgtgcaattcactcaagaanaattctctgttgatacaactctgataagagggatggt 320
QY 241 tagctcggttccctatgtttgattggtgtgtgtagagacttcagaanaatgctgtgtg 300
|||||
DB 321 tagctcggttccctatgtttgattggtgtgtgtagagacttcagaanaatgctgtgtg 380

RESULT 12
US-09-654-617-388605
Sequence 388605, Application US/09654617
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jindong
TITLE OF INVENTION: Annotated Plant Genes
FILE REFERENCE: 38-21(15097)D
CURRENT APPLICATION NUMBER: US/09/654,617
CURRENT FILING DATE: 2000-09-05
NUMBER OF SEQ ID NOS: 463173
SEQ ID NO 388605
LENGTH: 1056
TYPE: DNA
ORGANISM: Oryza sativa
US-09-654-617-388605

Query Match 74.6%; Score 223.8; DB 25; Length 1056;
Best Local Similarity 88.4%; Pred. No. 4.2e-62;
Matches 243; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 26 tgggattgatacctaattgctcaactagagacttgaactggaactactctgttaattgt 85
|||||
DB 1 tgggattgatacctaattgctcaactagagacttgaactggaactactctgttaattgt 60
QY 86 cattcttaagagcgagagacacactgaaccgcaatggtctactagaagtacagaagaac 145
|||||
DB 61 cattcttaagagcgagagacacactgaaccgcaatggtctactagaagtacagaagaac 120
QY 146 ctccagagatcggtggtgctgaattctcgaaacaggtgtgtgcaattactcaagaanaatc 205
|||||
DB 121 ctccagagatcggtggtgctgaattctcgaaacaggtgtgtgcaattactcaagaanaatc 180
QY 206 tctgttgataaactgctagagatagaggtgtgtagctggttctcctatgtttgattg 265
|||||
DB 181 tctgttgataaactgctagagatagaggtgtgtagctggttctcctatgtttgattg 240
QY 266 ggttggttgtagacttcagaanaatgctgtgtg 300
|||||
DB 241 ggttggttgtagacttcagaanaatgctgtgtg 275

RESULT 13
US-09-684-016-388605
Sequence 388605, Application US/09684016
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Liu, Jindong
TITLE OF INVENTION: Annotated Plant Genes
FILE REFERENCE: 38-21(15097)D
CURRENT APPLICATION NUMBER: US/09/684,016


```

; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 388605
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-684-016-388605
```

```

Query Match          74.6%; Score 223.8; DB 27; Length 1056;
Best Local Similarity 88.4%; Pred. No. 4.2e-62;
Matches 243; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
```

```

QY 26 tggagttatcatcaaatgtctcaactagtagacctgaacctgagcaactcttgaattgt 85
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 tggatcgcaccatcaaatgtctcaactagtagacagagctgtcactctcttgaattgt 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 86 catctcagaagcgagcgacactggaacccgcaatggtctactagaagttacagaagc 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 tattcttaagaagtgaagcacactggaacccggaatggtctcttgaagtgcacaaagc 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 146 ctcaagaagtgcgggctcgaactctcgaacaaagggtgttcgaattactcaagaataatc 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 ctccgagatgctggtcgtctctcgaacaaagggtgttcgaattactcaagaataatc 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 206 tctgttgatatacactcgtcgaataagaagaggtgttagctcgtttccctatgtttgatg 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 181 tctgttgatatacactcgtcgaataagaagaggtgttagctcgtttccctatgtttgatg 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 266 ggtcgtcgttagagactcagaataatgctgtctgtg 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 241 ggtcgtcgttagagactcagaataatgctgtctgtg 275
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 14
; Sequence 21745, Application US/09849529A
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen L.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51893)B
; CURRENT APPLICATION NUMBER: US/09/849,529A
; CURRENT FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: US 60/196,868
; PRIOR FILING DATE: 2000-05-09
; NUMBER OF SEQ ID NOS: 24076
; SEQ ID NO 21745
; LENGTH: 509
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3585-052-P1-K1-C6
US-09-849-529A-21745
```

```

Query Match          72.8%; Score 218.4; DB 32; Length 509;
Best Local Similarity 83.0%; Pred. No. 1.8e-60;
Matches 249; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
```

```

QY 1 gatttatgaacaacgcgtcctgctggtgatgcatcaatctgctcaactaggacctg 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 71 gatttatgataatacgtctcgtgcatgcatcaatctgctcaactaggacctg 130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 aactggcaactactcttgaattgtcaatttcaagaagcgagcgacactgaaccgcga 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 131 agttgctctactcttgaattgtcaatttcaagaagcgagcgacactgaaccgcga 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 atggtctactagaagttacagaagccttcagaagcgagcgagcgacactgaaccgcg 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

DB 191 atggtctactagaagttacagaagccttcagaagcgagcgagcgacactgaaccgcg 250
QY 181 gtcttcaattactcaagaanaattctctgttgataaacctgctagaatagaagatgt 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 251 gtgtgtctataacacaagaanaattctctgttgataaacctgctagaatagaagatgt 310
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 taagctggttcctctatgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttg 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 311 tagccaggttcctctatgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttg 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 15
US-60-196-868-21723
; Sequence 21723, Application US/60196868
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen L.
; APPLICANT: McCarter, David W.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WI
; FILE REFERENCE: 38-21(51893)A
; CURRENT APPLICATION NUMBER: US/60/196,868
; CURRENT FILING DATE: 2000-05-09
; NUMBER OF SEQ ID NOS: 24054
; SEQ ID NO 21723
; LENGTH: 509
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; OTHER INFORMATION: Clone ID: LIB3585-052-P1-K1-C6
US-60-196-868-21723
```

```

Query Match          72.8%; Score 218.4; DB 58; Length 509;
Best Local Similarity 83.0%; Pred. No. 1.8e-60;
Matches 249; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
```

```

QY 1 gatttatgaacaacgcgtcctgctggtgatgcatcaatctgctcaactaggacctg 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 71 gatttatgataatacgtctcgtgcatgcatcaatctgctcaactaggacctg 130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 aactggcaactactcttgaattgtcaatttcaagaagcgagcgacactgaaccgcga 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 131 agttgctctactcttgaattgtcaatttcaagaagcgagcgacactgaaccgcga 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 atggtctactagaagttacagaagccttcagaagatgcgggctgcaattctcgaacag 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 191 atggtctactagaagttacagaagccttcgaagatgcgggctgcaattctcgaacag 250
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 gttgtcaattactcaagaanaattctctgttgataaacctgctagaatagaagatgt 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 251 gtgtgtctataacacaagaanaattctctgttgataaacctgctagaatagaagatgt 310
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 taagctggttcctctatgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttg 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 311 tagccaggttcctctatgtttgtttgtttgtttgtttgtttgtttgtttgtttgtttg 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: July 3, 2002, 11:18:17
Job time: 34846 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2002, 08:11:53 ; Search time 185.98 seconds
(without alignments)
396.226 Million cell updates/sec

Title: US-09-300-482-619

Perfect score: 300
Sequence: 1 gattatgtgacaacacgat.....ttcagaatgtctgtgtg 300

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*

1: /cgn2_6/ptodata/2/1na/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/1na/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/1na/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/1na/PCITUS.COMB.seq:*
6: /cgn2_6/ptodata/2/1na/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	30.6	10.2	1051	3	US-08-961-083-125
2	29.8	9.9	1750	3	US-08-859-167-1
3	29.8	9.9	1750	3	US-09-109-273-1
4	29.8	9.9	1750	4	US-09-276-993-1
5	29.8	9.9	2040	4	US-09-069-023-33
6	29.8	9.9	2045	4	US-08-484-493-1
7	29.8	9.9	2297	1	US-08-484-493-1
8	29.8	9.9	2297	1	US-08-484-494-1
9	29.8	9.9	2297	2	US-08-345-212-1
10	29.8	9.9	2297	3	US-09-249-003-1
11	29.8	9.9	4428	1	US-08-484-493-6
12	29.8	9.9	4428	2	US-08-484-494-6
13	29.8	9.9	4428	2	US-08-345-212-6
14	29.8	9.9	4428	3	US-09-249-003-6
15	29.8	9.9	4546	4	US-09-134-218-2
16	29.8	9.9	1932	3	US-08-714-918-97
17	29.8	9.9	1932	4	US-09-265-315-97
18	29.8	9.9	1932	4	US-09-265-315-97
19	29.8	9.9	1932	4	US-09-265-315-97
20	29.8	9.9	15225	2	US-08-892-403A-2
21	29.8	9.9	1146	1	US-08-487-810-1
22	29.8	9.9	3906	2	US-08-469-537A-102
23	29.8	9.9	4165	1	US-08-442-248-1
24	29.8	9.9	4165	1	US-08-440-815-1
25	29.8	9.9	4165	4	US-08-486-448-1
26	29.8	9.9	351	4	US-09-660-877-17
27	29.8	9.9	1055	1	US-08-618-464-1

C 28	27.4	9.1	1055	3	US-09-107-615-1	Sequence 1, Appl
C 29	27.4	9.1	1055	5	PCT-US95-04896-1	Sequence 1, Appl
C 30	27.4	9.1	1881	4	US-09-553-498-1	Sequence 1, Appl
C 31	27.4	9.1	1881	4	US-09-553-498-3	Sequence 3, Appl
C 32	26.8	8.9	2441	1	US-08-332-312-1	Sequence 1, Appl
C 33	26.8	8.9	2624	1	US-08-032-382B-1	Sequence 1, Appl
C 34	26.8	8.9	5648	5	PCT-US96-03940-7	Sequence 7, Appl
C 35	26.8	8.9	5648	5	PCT-US96-03940-1	Sequence 1, Appl
C 36	26.6	8.9	1426	4	US-08-858-207A-78	Sequence 78, Appl
C 37	26.4	8.8	665	2	US-08-980-060-12	Sequence 12, Appl
C 38	26.4	8.8	665	4	US-09-307-185-12	Sequence 12, Appl
C 39	26.4	8.8	1236	4	US-09-445-472-2	Sequence 2, Appl
C 40	26.4	8.8	1566	4	US-08-894-818B-4	Sequence 4, Appl
C 41	26.4	8.8	1806	2	US-08-980-060-1	Sequence 1, Appl
C 42	26.4	8.8	1806	4	US-09-307-185-1	Sequence 1, Appl
C 43	26.4	8.8	1962	4	US-08-894-818B-34	Sequence 34, Appl
C 44	26.4	8.8	1962	4	US-09-445-472-15	Sequence 15, Appl
C 45	26.4	8.8	3089	1	US-07-921-796-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-961-083-125/c
; Sequence 125, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1051 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-083-125

Query Match 10.2%; Score 30.6; DB 3; Length 1051;
Best Local Similarity 53.8%; Pred. No. 0.78;
Matches 63; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 12 aaacacgacgcctgctggatgatcatcaaatgcccactagagactggaactgggaact 71
DB 677 AACCCACGTCCTCTTGGCAGTACGACAGGAGGTAATCTTCACATGCTTGGCGCAAGCTGAT 618

QY **72** acctcgttaattgcatcattctaagagcgagacaccctgaacccgcgatgctca 128
 | ||| ||| ||||| ||| || |||||
Db 617 AATCTCGTAGTGTGCTTTTGAGGACGCACATCCACTTCATAACCTAAGGCTCTA 561

RESULT 2
US-09-050

```

US-08-859-167-1
/ Sequence 1, Application US/08859167
/ Patent No. 6037461
/ GENERAL INFORMATION:
/ APPLICANT: Alnemri, Emad S.
/ APPLICANT: Fernandez-Alnemri, Teresa
/ TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
/ TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
/ TITLE OF INVENTION: OF MAKING THE SAME
/ NUMBER OF SEQUENCES: 17
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6037461rlis
/ STREET: One Liberty Place, 46th floor
/ CITY: Philadelphia
/ STATE: PA
/ COUNTRY: USA
/ ZIP: 19103
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: WINDOWS
/ SOFTWARE: WordPerfect
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/859,167
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Deluca, Mark
/ REGISTRATION NUMBER: 33,229
/ REFERENCE/DOCKET NUMBER: TJU-
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (215) 568-3100
/ TELEFAX: (215) 568-3439
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1750 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: both
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 413..1750
/ US-08-859-167-1

```

```

? Patent No. 6063760
? GENERAL INFORMATION:
? APPLICANT: Alnemri, Emad S.
? APPLICANT: Fernandez-Alnemri, Teresa
? TITLE OF INVENTION: FAD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
? TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
? TITLE OF INVENTION: OF MAKING THE SAME
? NUMBER OF SEQUENCES: 17
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6063760-rs
? STREET: One liberty place, 46th floor
? CITY: Philadelphia
? STATE: PA
? COUNTRY: USA
? ZIP: 19103
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: WINDOWS
? SOFTWARE: Wordperfect
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/109,273
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/859,167
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Deluca, Mark
? REGISTRATION NUMBER: 33,229
? REFERENCE/DOCKET NUMBER: TCU-
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (215) 568-3100
? TELEFAX: (215) 568-3439
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1750 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: both
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 413..1750
? US-09-109-273-1

```

QY 174

```
Db      485 TGCCGGATGTTCCTAT 501
          |  |||||
RESULT   3
US-09-109-273-1
; Sequence 1, Application US/09109273
```

RESULT 4
NE-09-276-

US 9,276,993-1
Sequence 1, Application US/09276993
Patent No. 6207801
GENERAL INFORMATION:
APPLICANT: Alnemri, Emad S.
APPLICANT: Fernandez-Alnemri, Teresa
TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS


```
RESULT 9
US-08-345-212-1
; Sequence 1, Application US/08345212
; Patent No. 5932211
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielicki, Julie
; APPLICANT: Clements, Peter R
; APPLICANT: Hopwood, John J
; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
; NUMBER OF SEQUENCES: 15
; IDURONATE 2-SULFATASE
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/345,212
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 991,973
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8416Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2297 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 125..1774
US-08-345-212-1

Query Match          9.7%; Score 29; DB 2; Length 2297;
Best Local Similarity 53.0%; Pred. No. 3.8;
Matches 62; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 121 atggtctactagaagtaacgaagaagccttcagagatcggggctgcgaatttcgaacaagg 180
DB 1505 ATTGCTATATAGCCATATCCCGGCTTCAGACATCCCTCAGATGGAATTCTGACAAGCCG 1564

QY 181 gtgtgcaattactcaagaanaattctctgttgataatacaactgctagaataagagat 237
DB 1565 AGTTTAAAGATATTAAGATCATGGCTATTCCATACGACCATATAGACTATAGGTAT 1621

RESULT 10
US-09-249-003-1
; Sequence 1, Application US/09249003
; Patent No. 6153188
; GENERAL INFORMATION:
```

```
APPLICANT: Wilson, Peter J
APPLICANT: Morris, Charles P
APPLICANT: Anson, Donald S
APPLICANT: Occhiodoro, Teresa
APPLICANT: Bielicki, Julie
APPLICANT: Clements, Peter R
APPLICANT: Hopwood, John J
TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
NUMBER OF SEQUENCES: 15
IDURONATE 2-SULFATASE
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/249,003
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/991,973
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8416Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2297 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 125..1774
US-09-249-003-1

Query Match          9.7%; Score 29; DB 3; Length 2297;
Best Local Similarity 53.0%; Pred. No. 3.8;
Matches 62; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 121 atggtctactagaagtaacgaagaagccttcagagatcggggctgcgaatttcgaacaagg 180
DB 1505 ATTGCTATATAGCCATATCCCGGCTTCAGACATCCCTCAGATGGAATTCTGACAAGCCG 1564

QY 181 gtgtgcaattactcaagaanaattctctgttgataatacaactgctagaataagagat 237
DB 1565 AGTTTAAAGATATTAAGATCATGGCTATTCCATACGACCATATAGACTATAGGTAT 1621

RESULT 11
US-08-484-493-6
; Sequence 6, Application US/08484493
; Patent No. 5728381
; GENERAL INFORMATION:
; APPLICANT: Wilson, Peter J
; APPLICANT: Morris, Charles P
; APPLICANT: Anson, Donald S
; APPLICANT: Occhiodoro, Teresa
; APPLICANT: Bielicki, Julie
```


1	APPLICANT:	Clements, Peter R
2	APPLICANT:	Hopwood, John J
3	TITLE OF INVENTION:	GLYCOSYLATION VARIANTS OF
4	TITLE OF INVENTION:	IDRONATE 2-SULFATASE
5	NUMBER OF SEQUENCES:	15
6	CORRESPONDENCE ADDRESS:	
7	ADDRESSEE:	Scully, Scott, Murphy & Presser
8	STREET:	400 Garden City Plaza
9	CITY:	Garden City
10	STATE:	New York
11	COUNTRY:	USA
12	ZIP:	11530
13	COMPUTER READABLE FORM:	
14	MEDIUM TYPE:	Floppy disk
15	COMPUTER:	IBM PC compatible
16	OPERATING SYSTEM:	PC-DOS/MS-DOS
17	SOFTWARE:	Patentin Release #1.0, Version #1.25
18	CURRENT APPLICATION DATA:	
19	APPLICATION NUMBER:	US/08/484,493
20	FILING DATE:	07-JUN-1995
21	CLASSIFICATION:	424
22	PRIOR APPLICATION DATA:	
23	APPLICATION NUMBER:	991,973
24	FILING DATE:	17-DEC-1992
25	ATTORNEY/AGENT INFORMATION:	
26	NAME:	DiGiglio, Frank S
27	REGISTRATION NUMBER:	31,346
28	REFERENCE/DOCKET NUMBER:	84162
29	TELECOMMUNICATION INFORMATION:	
30	TELEPHONE:	516-742-4343
31	TELEFAX:	516-742-4366
32	TELEX:	230 901 SANS UR
33	INFORMATION FOR SEQ ID NO:	6:
34	SEQUENCE CHARACTERISTICS:	
35	LENGTH:	4428 base pairs
36	TYPE:	nucleic acid
37	STRANDEDNESS:	double
38	TOPOLOGY:	linear
39	MOLECULE TYPE:	DNA (genomic)
40	FEATURE:	
41	NAME/KEY:	CDS
42	LOCATION:	332..434
43	FEATURE:	
44	NAME/KEY:	Intron
45	LOCATION:	536..537
46	FEATURE:	
47	NAME/KEY:	CDS
48	LOCATION:	693..829
49	FEATURE:	
50	NAME/KEY:	Intron
51	LOCATION:	962..963
52	FEATURE:	
53	NAME/KEY:	CDS
54	LOCATION:	1044..1221
55	FEATURE:	
56	NAME/KEY:	Intron
57	LOCATION:	1350..1351
58	FEATURE:	
59	NAME/KEY:	CDS
60	LOCATION:	1480..1569
61	FEATURE:	
62	NAME/KEY:	Intron
63	LOCATION:	1716..1717
64	FEATURE:	
65	NAME/KEY:	CDS
66	LOCATION:	1841..2041
67	FEATURE:	
68	NAME/KEY:	Intron
69	LOCATION:	2206..2207
70	FEATURE:	
71	NAME/KEY:	CDS
72	LOCATION:	2294..2464
73	FEATURE:	

1	NAME/KEY:	Intron
2	LOCATION:	2585..2588
3	FEATURE:	
4	NAME/KEY:	CDS
5	LOCATION:	2684..2810
6	FEATURE:	
7	NAME/KEY:	Intron
8	LOCATION:	2904..2905
9	FEATURE:	
10	NAME/KEY:	CDS
11	LOCATION:	3033..3206
12	FEATURE:	
13	NAME/KEY:	Intron
14	LOCATION:	3308..3309
15	FEATURE:	
16	NAME/KEY:	CDS
17	LOCATION:	3435..3908
18	FEATURE:	
19	NAME/KEY:	Intron
20	LOCATION:	4804..4933
21	FEATURE:	
22	NAME/KEY:	CDS
23	LOCATION:	4933..4936
24	FEATURE:	

```
Query Match      9.7%; Score 29; DB 1; Length 4428;
Best Local Similarity 53.0%; Pred. No. 5.1;
Matches 62; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
```

QY 121 atggctcaccacagaagacagaagaagccttcagagatgcgagggctgcgaattctcgaaacag 180
 Db 3636 ATGGCTCATAGCCAGTATCCCGCCCTTCAGACATCCCTCAATGAAATTCCTACACGCC 36395
 QY 181 gtgttgaaatctactcaagaanaattctctgttgataacacgtctgaataagagat 237
 Db 3696 AGTTTAAAGATTTAAGATCATGAGGCTATTCATTCACGACCATTAAGACTTAAGTAT 3752

RESULT 12

US-08-484-494-6
Sequence 6, Application US/08484494
Patent No. 5708330

GENERAL INFORMATION:
 APPLICANT: Wilson, Peter J
 APPLICANT: Morris, Charles P
 APPLICANT: Anson, Donald S
 APPLICANT: Occhiodoro, Teresa
 APPLICANT: Bielikli, Julie
 APPLICANT: Clements, Peter R
 APPLICANT: Hopwood, John J
 TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
 TITLE OF INVENTION: IDURONATE 2-SULFATASE
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:

ADDRESSEE: scully, scott, murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
CARD: 11/100

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,494
FILING DATE: 07-JUN-1995
CLASSIFICATION:

APPLICATION NUMBER: 991,973
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digislo, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8416
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343

1	TELEFAX: 516-742-4366
2	TELEX: 230 901 SANS UR
3	INFORMATION FOR SEQ ID NO: 6:
4	SEQUENCE CHARACTERISTICS:
5	LENGTH: 4428 base pairs
6	type: nucleic acid
7	STRANDEDNESS: double
8	TOPOLOGY: linear
9	MOLECULE TYPE: DNA (genomic)
10	FEATURE:
11	NAME/KEY: CDS
12	LOCATION: 332..434
13	FEATURE:
14	NAME/KEY: Intron
15	LOCATION: 536..537
16	FEATURE:
17	NAME/KEY: CDS
18	LOCATION: 693..829
19	FEATURE:
20	NAME/KEY: Intron
21	LOCATION: 962..963
22	FEATURE:
23	NAME/KEY: CDS
24	LOCATION: 1044..1221
25	FEATURE:
26	NAME/KEY: Intron
27	LOCATION: 1350..1351
28	FEATURE:
29	NAME/KEY: CDS
30	LOCATION: 1480..1569
31	FEATURE:
32	NAME/KEY: Intron
33	LOCATION: 1716..1717
34	FEATURE:
35	NAME/KEY: CDS
36	LOCATION: 1841..2041
37	FEATURE:
38	NAME/KEY: Intron
39	LOCATION: 2206..2207
40	FEATURE:
41	NAME/KEY: CDS
42	LOCATION: 2294..2464
43	FEATURE:
44	NAME/KEY: Intron
45	LOCATION: 2585..2586
46	FEATURE:
47	NAME/KEY: CDS
48	LOCATION: 2684..2810
49	FEATURE:
50	NAME/KEY: Intron
51	LOCATION: 2994..2905
52	FEATURE:
53	NAME/KEY: CDS
54	LOCATION: 3033..3206
55	FEATURE:
56	NAME/KEY: Intron
57	LOCATION: 3308..3309
58	FEATURE:
59	NAME/KEY: CDS
60	LOCATION: 3435..3908
61	LOCATION: 484-494-6

Query Match	9.7%	Score 29	DB 1	Length 4428
Best Local Similarity	53.0%	Pred. No. 5.1		
Matches 62	Conservative 0	Mismatches 55	Indels 0	Gaps 0
QY 121	atggtctactagaagtaacagaagaagccttcacagatgcg99ggtcgtacatctcgaacag9	180		
Db 3636	ATTGGCTATTAGCCACTAATCCCGCGCTTCAGACATCCCTCACTGGAGATTCTGACAAAGCG	3695		
QY 181	gtgtgtgcattactactcaagaagaattctctgttgcataacacatcgtctgaatagagag9	237		

```

Db      3696  AGTTTAAGATATTAAGATCATGGCGCTATTCATACGACCATGAGCTATAGGTAT 3752
RESULT 13
US-08-345-212-6
: Sequence 6, Application US/08345212
: Patent No. 5932211
GENERAL INFORMATION:
APPLICANT: Wilson, Peter J
APPLICANT: Morris, Charles P
APPLICANT: Anson, Donald S
APPLICANT: Occhiodoro, Teresa
APPLICANT: Bielecki, Julie
APPLICANT: Clements, Peter R
APPLICANT: Hopwood, John J
TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
TITLE OF INVENTION: IDUCONATE 2-SULFATASE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
City: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/345,212
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 991,973
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8416Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 4428 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 332..434
FEATURE:
NAME/KEY: Intron
LOCATION: 536..537
FEATURE:
NAME/KEY: CDS
LOCATION: 693..829
FEATURE:
NAME/KEY: Intron
LOCATION: 962..963
FEATURE:
NAME/KEY: CDS
LOCATION: 1044..1221
FEATURE:
NAME/KEY: Intron
LOCATION: 1350..1351
FEATURE:
NAME/KEY: CDS
LOCATION: 1480..1569

```


FEATURE:
NAME/KEY: Intron
LOCATION: 1716..1717
FEATURE:
NAME/KEY: CDS
LOCATION: 1841..2041
FEATURE:
NAME/KEY: Intron
LOCATION: 2206..2207
FEATURE:
NAME/KEY: CDS
LOCATION: 2294..2464
FEATURE:
NAME/KEY: Intron
LOCATION: 2585..2586
FEATURE:
NAME/KEY: CDS
LOCATION: 2684..2810
FEATURE:
NAME/KEY: Intron
LOCATION: 2904..2905
FEATURE:
NAME/KEY: CDS
LOCATION: 3033..3206
FEATURE:
NAME/KEY: Intron
LOCATION: 3308..3309
FEATURE:
NAME/KEY: CDS
LOCATION: 3435..3908
US-08-345-212-6

Query Match
Best Local Similarity 9.7%; Score 29; DB 2; Length 4428;
Matches 62; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 121 atgttactaagatgacagaagccttcagagatcg9ggtgcaatctcgaacagg 180
Db 3636 attgcctatagccagatgcccggccttcacagatccctcagtggaatttcgacacg 3635
QY 181 gtgttgcaatctcgaagaatctctgtgtgataacacgcgcgaatagaggat 237
Db 3696 agtttaaaagatataaagatcttggtctatttcacacgacacatagactatagctat 3752

RESULT 14
US-09-249-003-6
Sequence 6, Application US/09249003
Patent No. 615318
GENERAL INFORMATION:
APPLICANT: Wilson, Peter J
APPLICANT: Morris, Charles P
APPLICANT: Anson, Donald S
APPLICANT: Occhiodoro, Teresa
APPLICANT: Bielicki, Julie
APPLICANT: Hopwood, John R
TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/249,003
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/991,973
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digilio, Frank S
REGISTRATION NUMBER: 31,346
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 4428 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 332..434
FEATURE:
NAME/KEY: Intron
LOCATION: 536..537
FEATURE:
NAME/KEY: CDS
LOCATION: 693..829
FEATURE:
NAME/KEY: Intron
LOCATION: 962..963
FEATURE:
NAME/KEY: CDS
LOCATION: 1044..1221
FEATURE:
NAME/KEY: Intron
LOCATION: 1350..1351
FEATURE:
NAME/KEY: CDS
LOCATION: 1480..1569
FEATURE:
NAME/KEY: Intron
LOCATION: 1716..1717
FEATURE:
NAME/KEY: CDS
LOCATION: 1841..2041
FEATURE:
NAME/KEY: Intron
LOCATION: 2206..2207
FEATURE:
NAME/KEY: CDS
LOCATION: 2294..2464
FEATURE:
NAME/KEY: Intron
LOCATION: 2585..2586
FEATURE:
NAME/KEY: CDS
LOCATION: 2684..2810
FEATURE:
NAME/KEY: Intron
LOCATION: 2904..2905
FEATURE:
NAME/KEY: CDS
LOCATION: 3033..3206
FEATURE:
NAME/KEY: Intron
LOCATION: 3308..3309
FEATURE:
NAME/KEY: CDS
LOCATION: 3435..3908

US-09-249-003-6

Query Match	9.7%	Score 29;	DB 3;	Length 4428;
Best Local Similarity	53.0%;	Pred. No. 5.1;		
Matches	62;	Conservative	0;	Mismatches 55;
			Indels	0;
			Gaps	0;

Qy 121 atgctgtactatgaagttacagaagaaccttcagatgcgggctgtcaattctcgaaacagg 180
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3636 ATTGCTATTAGCCAGTATCCCCGCCCTTCAGACATCCTCACTGAATTCTGCACAAGCG 3695

Ry 181 ggtctgcattactcaagaataatcctctgttggataacaetgctaataagaggat 237
|| || | ||| || ||| ||| ||| ||| ||| |||
Db 3696 AGTTTAAAGATATAAGACATGCGGTATTCCATACGCACCATAGACTATAGCTAT 3752

RESULT 15
 US-09-134-218-2/c
 : Sequence 2, Application US/09134218A
 : Patent No. 6312926
 : GENERAL INFORMATION:
 : APPLICANT: Shaktin, Aaron J.
 : APPLICANT: Pillutla, Renuka
 : APPLICANT: Reinberg, Danny
 : APPLICANT: Yu, Zheng
 : APPLICANT: Moldonado, Edito
 : TITLE OF INVENTION: MRNA CAPPING ENZYMES AND USES THEREOF
 : FILE REFERENCE: 601-1-079 ss
 : CURRENT APPLICATION NUMBER: US/09/134,218A
 : CURRENT FILING DATE: 1998-08-14
 : NUMBER OF SEQ ID NOS: 19
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 2
 : LENGTH: 4546
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : US-09-134-218-2

Query Match	9.48;	Score 28.2;	DB 4;	Length 4546;
Best Local Similarity	48.48;	Pred. No. 9.4;		
Matches 78;	Conservative 0;	Mismatches 83;	Indels 0;	Gaps 0;

Oy 134 agtcatagaagcccttcagayattgcggcgctgcaatctcgtgaacagagtgtttgcaattac 193
 ||| ||| | ||||| ||| ||| ||
Db 3139 AGTTCACAGAGTGCTACACTTTGCTCCTGCATTTATCTACAGCAGTGGAGAGCATTTTC 3080

```

QY 194 tcagaanaattctctgttgataaacactgcagaaacagaggaatggttagctcgatttc 253
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 3079 TGGTGCACGACATTCTATTCCATGACATTACTCTTAATCAAGAAGCTTAGCTGGCATTCC 3020

```

Qy	254	tatgcttgattggttggtgtgtaggaacttcgaatgtctg	294
		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
Db	3019	CTTTCGTCTTTCTTCGAGAGAAGAACACACATCGTTTG	2979

Search completed: July 3, 2002, 08:11:57
Job time: 25882 sec

THIS PAGE BLANK (USPTO)